

UCUENCA

Facultad de Ingeniería Carrera de Ingeniería de Sistemas

Desarrollo de un XBlock en Open edX para apoyar las Analíticas de Aprendizaje

Trabajo de titulación previo a la obtención del título de Ingeniero de Sistemas.

Autores:

Jonnathan Henry Campoberde Ávila

CI: 0105476097

Correo electrónico: jonnathan.campoberde@gmail.com

Miguel Ángel Macías Narváez

CI: 0706422250

Correo electrónico: mangel.maciasn@gmail.com

Director:

Ing. Jorge Javier Maldonado Mahauad, PhD.

CI: 1102959051

Cuenca, Ecuador

26-julio-2022

Resumen

Hoy en día, cuando los usuarios utilizan los sistemas de información, generan una gran cantidad de datos que dejan un rastro como resultado de su interacción, por ejemplo, cuando los estudiantes acceden a materiales educativos en Entornos Virtuales de Aprendizaje (VLE – Virtual Learning Environments). Esto se ha convertido en lo que ahora conocemos como Cursos Masivos Abiertos en Línea (MOOC – por sus siglas en inglés), que tienen millones de estudiantes registrados. Debido a la gran cantidad de datos que se generan dentro de los MOOC, las analíticas de aprendizaje (LA – Learning Analytics) ha surgido como una alternativa para mejorar los procesos de enseñanza y aprendizaje a través del análisis de datos. Open edX en un intento de incorporar LA en su plataforma de desarrollo Insights, que proporciona visualizaciones muy simples. De igual manera, se han sumado otros proyectos que con el paso del tiempo han quedado obsoletos o no brindan el apoyo suficiente para mejorar el proceso de aprendizaje de los estudiantes.

Debido a lo anterior, este trabajo de titulación propone el diseño, desarrollo y evaluación de un dashboard de analíticas de aprendizaje para la plataforma Open edX. La herramienta incorporará indicadores de éxito académico en sus visualizaciones para mejorar el proceso de aprendizaje dentro de la plataforma. Con el propósito de desarrollar la herramienta de LA, se realizó un análisis exploratorio del comportamiento de los estudiantes para determinar variables y secuencias de aprendizaje, se diseñaron visualizaciones para ser incorporadas dentro del dashboard de docentes y estudiantes. Para llevar a cabo el desarrollo del componente de LA, denominado XLEA (XBlock for LEarning Analytics), se empleó la metodología LATUX (Learning Awareness Tool – User eXperience), la cual consta de cinco etapas agrupadas en dos enfoques.

Finalmente, luego de haber evaluado la herramienta XLEA con el cuestionario Evaluation Framework for Learning Analytics (EFLA) y con el cuestionario de experiencia de usuario UEQ, por estudiantes y maestros de la Universidad de Cuenca, se logró evidenciar un alto grado de aceptación y conformidad por parte de los participantes al usar XLEA. La mayoría de participantes se enfocaron en aquellas visualizaciones que permiten tener una comprensión del comportamiento histórico por semanas de las diferentes actividades que desempeñan dentro del curso, por ejemplo, videos, lecturas y problemas.

Palabras clave: Learning analytics. Open edX. Dashboard. Éxito académico. XBlock.

Abstract

Nowadays, when users use information systems, they generate a large amount of data that leaves a trace as a result of their interaction, for example, when students access educational materials in Virtual Learning Environments (VLE). This has evolved into what we now know as Massive Open Online Courses (MOOC), which have millions of students registered. Due to the large amount of data generated within MOOC, Learning Analytics (LA) has emerged as an alternative to improve teaching and learning processes through data analysis. Open edX in an attempt to incorporate LA into its Insights development platform, which provides very simple visualizations. Similarly, other projects have been added that have become obsolete over time or do not provide sufficient support to improve the student learning process.

Due to the above, this degree work proposes the design, development and evaluation of a learning analytics dashboard for the Open edX platform. The tool will incorporate indicators of academic success in its visualizations to improve the learning process within the platform. With the purpose of developing the LA tool, an exploratory analysis of student behavior was carried out to determine variables and learning sequences, visualizations were designed to be incorporated into the dashboard of teachers and students. To carry out the development of the LA component, called XLEA (XBlock for LEarning Analytics), the LATUX (Learning Awareness Tool – User eXperience) methodology was used, which consists of five stages grouped into two approaches.

Finally, after having evaluated the XLEA tool with the Evaluation Framework for Learning Analytics (EFLA) questionnaire and with the UEQ user experience questionnaire, by students and teachers of the University of Cuenca, a high degree of acceptance and conformity was evidenced. by participants when using XLEA. Most of the participants focused on those visualizations that allow an understanding of the historical behavior for weeks of the different activities that they carry out within the course, for example, videos, readings and problems.

Keywords: Learning analytics. Open edX. Dashboard. Academic success. XBlock.

UCUENCA

3.3.	Estructura de los cursos.....	45
3.4.	Gestión de los datos	46
3.4.1.	Base de datos.....	47
3.4.2.	Tracking Logs.....	48
3.5.	Conclusiones	51
Capítulo 4: Análisis exploratorio del comportamiento de los estudiantes		52
4.1.	Introducción	53
4.2.	Análisis exploratorio de los datos	53
4.2.1.	Análisis de requerimientos.....	53
4.2.2.	Selección de tablas importantes de la base de datos	55
4.3.	Identificación de variables para el mejoramiento del éxito estudiantil .	56
4.3.1.	Identificación de indicadores.....	57
4.3.2.	Análisis y selección de los indicadores	60
4.4.	Análisis del comportamiento de los estudiantes	62
4.4.1.	Extracción de los datos de la plataforma	62
4.4.2.	Herramienta empleada para Minería de Procesos.....	68
4.4.3.	Estrategias de aprendizaje e Identificación de patrones del comportamiento dentro del curso	71
4.5.	Conclusiones	84
Capítulo 5: Diseño de las visualizaciones		86
5.1.	Introducción	87
5.2.	Revisión y selección de visualizaciones existentes	87
5.3.	Criterios para la selección de las visualizaciones.....	91

UCUENCA

5.4. Diseño de las visualizaciones	92
5.5. Conclusiones	94
Capítulo 6: Desarrollo del Xblock	96
6.1. Introducción	97
6.2. Implementación de las visualizaciones dentro de Open edX.....	97
6.2.1. Arquitectura de un XBlock	97
6.2.2. Funcionalidades	99
6.2.3. Desarrollo del XBlock	99
6.2.4. Visualizaciones del XBlock desarrollado en este trabajo de titulación	103
6.3. Instalación del XBlock en un ambiente de producción.....	108
6.4. Ejecución del XBlock en un ambiente de producción	113
6.5. Evaluación del componente de visualizaciones.....	117
6.5.1. Selección de los involucrados en la evaluación	119
6.5.2. Aplicación del cuestionario a los evaluadores.....	119
6.6. Resultados de la evaluación a estudiantes empleando EFLA	121
6.7. Resultados de la evaluación a docentes usando EFLA.....	127
6.8. Resultados de la evaluación de experiencia de usuario	133
6.9. Conclusiones	139
Capítulo 7: Conclusiones, Recomendaciones y Trabajos futuros	141
7.1. Introducción	142
7.2. Conclusiones	142
7.3. Recomendaciones	144

UCUENCA

7.4. Trabajos Futuros.....	145
Referencias	146
Anexos	159

Índice de Figuras

Figura 1.	Pasos de la metodología adaptada LATUX a emplear para el desarrollo del trabajo de titulación.....	30
Figura 2.	Características de los MOOC.	35
Figura 3.	Crecimiento de los MOOC a través de los años.	37
Figura 4.	Modelo de los patrones de aprendizaje.	38
Figura 5.	Elementos de la Minería de Procesos en un contexto educativo.	39
Figura 6.	Breve descripción de la arquitectura de Open edX.	44
Figura 7.	Estructura de un curso en Open edX.....	46
Figura 8.	Base de datos en la arquitectura de Open edX.....	48
Figura 9.	Descubrimiento automatizado de procesos.	68
Figura 10.	Animación del mapa de procesos.....	69
Figura 11.	Estadísticas detalladas de la herramienta.....	70
Figura 12.	Vista de casos en Disco.	71
Figura 13.	Etapas para generar el modelo de procesos.	71
Figura 14.	Distribución de las variables para caracterizar los grupos del cluster.	75
Figura 15.	Método del codo para determinar el número ideal de grupos.	76
Figura 16.	Método de la silueta para determinar el número ideal de grupos.	77
Figura 17.	Visualización obtenida empleando el método de reducción de dimensionalidad PCA.....	78
Figura 18.	Visualización obtenida empleando el método de reducción de dimensionalidad t-SNE.....	78
Figura 19.	Modelo de procesos para el Clúster 0.	80
Figura 20.	Modelo de procesos para el Clúster 1	81

UCUENCA

Figura 21. Modelo de procesos para el Clúster 2.	82
Figura 22. Dashboard de edX Insights para medir el nivel de compromiso de los estudiantes dentro de un curso.....	88
Figura 23. Visualización del resumen de un curso.....	89
Figura 24. Visualización del tiempo invertido en cada capítulo.	89
Figura 25. Visualización de la cantidad de accesos de cada capítulo.....	90
Figura 26. Visualización del progreso de videos y problemas.....	90
Figura 27. (a) Menu principal de OXALIC. (b) Resumen del número total de interacciones con los vídeos. (c) Número de interacciones para cada vídeo. (d) Distribución de eventos relacionados a la reproducción de los videos. (e) Tiempo en la plataforma, inscripciones, sesiones y duración.	91
Figura 28. Prototipo del dashboard de la vista del estudiante.....	93
Figura 29. Prototipo del dashboard de la vista del docente.....	94
Figura 30. Arquitectura propuesta para el XBlock de LA.	99
Figura 31. Instalación del XBlock SDK.	100
Figura 32. Comando para crear un XBlock usando el XBlock SDK.	101
Figura 33. Estructura de las carpeta y archivos de un XBlock.	102
Figura 34. Instalación y ejecución del XBlock dentro del SDK Server.....	103
Figura 35. Rendimiento general por semanas. Vista del estudiante.	104
Figura 36. Tiempo invertido en el curso por semanas. Vista del estudiante.	104
Figura 37. Recursos completados a lo largo del curso. Vista del estudiante.....	105
Figura 38. Recursos completados por semanas de todos los estudiantes del curso. Vista del docente.	106
Figura 39. Interacciones con videos de todos los estudiantes del curso. Vista del docente.	106

UCUENCA

Figura 40. Sesiones promedio por semana y recursos accedidos en promedio por semana. Vista del docente.	107
Figura 41. Estadísticas de los indicadores de cada estudiante del curso. Vista del docente.	108
Figura 42. Migración del código de XLEA al servidor de producción.	109
Figura 43. Otorgar permisos al usuario “edxapp” sobre el directorio del componente de LA.	110
Figura 44. Instalación de las dependencias del componente de LA.	111
Figura 45. Instalación del componente de LA en Open edX.	111
Figura 46. Actualización del LMS en Open edX.	112
Figura 47. Reinicio de todos los servicios de Open edX.	112
Figura 48. Inicio de sesión en el dashboard LMS del Open edX.	113
Figura 49. Selección del curso donde se va habilitar el XBlock.	114
Figura 50. Seleccionar configuración avanzada del curso.	114
Figura 51. Habilitación del XBlock desarrollado en este trabajo de titulación.	115
Figura 52. Creación de la unidad donde será mostrado el XBlock dentro del curso.	115
Figura 53. Agregación de nuevo componente avanzado dentro del curso.	116
Figura 54. Selección del XBlock a mostrar dentro de la unidad del curso.	116
Figura 55. Previsualización del dashboard de la vista del estudiante dentro del curso.	117
Figura 56. Dimensiones de evaluación de herramientas de LA.	118
Figura 57. Cuestionario para la evaluación de herramientas de LA, estudiante y docente.	120
Figura 58. Respuestas a la primera pregunta del EFLA para estudiantes.	122

UCUENCA

Figura 59.	Repuestas a la segunda pregunta del EFLA para estudiantes.	123
Figura 60.	Repuestas a la tercera pregunta del EFLA para estudiantes.	124
Figura 61.	Repuestas a la cuarta pregunta del EFLA para estudiantes.	124
Figura 62.	Repuestas a la quinta pregunta del EFLA para estudiantes.	125
Figura 63.	Repuestas a la sexta pregunta del EFLA para estudiantes.	125
Figura 64.	Repuestas a la séptima pregunta del EFLA para estudiantes.	126
Figura 65.	Repuestas a la octava pregunta del EFLA para estudiantes.	127
Figura 66.	Repuestas a la primera pregunta del EFLA para profesores.	128
Figura 67.	Repuestas a la segunda pregunta del EFLA para profesores.	129
Figura 68.	Repuestas a la tercera pregunta del EFLA para profesores.	130
Figura 69.	Repuestas a la cuarta pregunta del EFLA para profesores.	130
Figura 70.	Repuestas a la quinta pregunta del EFLA para profesores.	131
Figura 71.	Repuestas a la sexta pregunta del EFLA para profesores.	131
Figura 72.	Repuestas a la séptima pregunta del EFLA para profesores.	132
Figura 73.	Repuestas a la octava pregunta del EFLA para profesores.	133
Figura 74.	Repuestas a la primera pregunta del UEQ.	134
Figura 75.	Repuestas a la segunda pregunta del UEQ.	135
Figura 76.	Repuestas a la tercera pregunta del UEQ.	136
Figura 77.	Repuestas a la cuarta pregunta del UEQ.	137
Figura 78.	Repuestas a la quinta pregunta del UEQ.	138
Figura 79.	Repuestas a la sexta pregunta del UEQ.	139

Índice de Tablas

Tabla 1.	Vista general de las herramientas de visualización que existen para Open edX.	27
Tabla 2.	Descripción de los campos que son registrados en el archivo tracking log.	49
Tabla 3.	Descripción de las tablas seleccionadas que contienen información relevante acerca de las interacciones de los estudiantes en la plataforma.....	56
Tabla 4.	Indicadores relacionados al éxito académico en MOOC.....	58
Tabla 5.	Indicadores seleccionados para implementarse en el dashboard de LA.	60
Tabla 6.	Eventos registrados por los estudiantes cuando interactúan con la interacción con los videos.	63
Tabla 7.	Eventos registrados por los estudiantes con las evaluaciones.....	64
Tabla 8.	Eventos registrados por los estudiantes en los foros.	65
Tabla 9.	Eventos registrados con la interacción de recursos de lecturas... ..	65
Tabla 10.	Tabla creada para almacenar la información procesada de los trackings logs.	67
Tabla 11.	Análisis usando la media de los clusters obtenidos.	79
Tabla 12.	Dimensiones de evaluación de herramientas de LA.	118
Tabla 13.	Estructura del cuestionario UEQ.....	121
Tabla 14.	Resultados de aplicar el cuestionario EFLA en los estudiantes. .	127
Tabla 15.	Resultados de aplicar el cuestionario EFLA en los docentes.	133

Índice de Anexos

Anexo 1: Evento guardado en el archivo tracking logs.....	159
Anexo 2: Descripción de las tablas que contienen información acerca de los estudiantes dentro de la plataforma.	160
Anexo 3: Scripts y comandos para migrar los datos	162
Anexo 4: Scripts que analiza los eventos generados por los estudiantes ...	163
Anexo 5: Enlace al repositorio de Github con los Scripts que analiza las actividades realizadas y obtiene estadísticas en forma de sesiones.	168
Anexo 6: Artículo publicado en la conferencia LACLO 2021 con doi: 10.1109/LACLO54177.2021.00088.....	168

Cláusula de licencia y autorización para publicación en el Repositorio Institucional

Jonnathan Henry Campoberde Ávila en calidad de autor/a y titular de los derechos morales y patrimoniales del trabajo de titulación "Desarrollo de un XBlock en Open edX para apoyar las Analíticas de Aprendizaje", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 26 de julio de 2022



Jonnathan Henry Campoberde Ávila

C.I: 0105476097

Cláusula de Propiedad Intelectual

Jonnathan Henry Campoberde Ávila, autor/a del trabajo de titulación “Desarrollo de un XBlock en Open edX para apoyar las Analíticas de Aprendizaje”, certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor/a.

Cuenca, 26 de julio de 2022



Jonnathan Henry Campoberde Ávila

C.I: 0105476097

Cláusula de licencia y autorización para publicación en el Repositorio Institucional

Miguel Ángel Macías Narváez en calidad de autor/a y titular de los derechos morales y patrimoniales del trabajo de titulación "Desarrollo de un XBlock en Open edX para apoyar las Analíticas de Aprendizaje", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 26 de julio de 2022



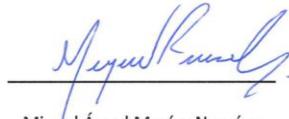
Miguel Ángel Macías Narváez

C.I: 0706422250

Cláusula de Propiedad Intelectual

Miguel Ángel Macías Narváez, autor/a del trabajo de titulación "Desarrollo de un XBlock en Open edX para apoyar las Analíticas de Aprendizaje", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor/a.

Cuenca, 26 de julio de 2022



Miguel Ángel Macías Narváez

C.I.: 0706422250

Listado de Abreviaturas

- **LA.** – *Learning Analytics*: Analíticas de Aprendizaje.
- **VLE.** – *Virtual Learning Environment*: Ambiente de Aprendizaje Virtual.
- **LMS.** – *Learning Management System*: Sistema de Gestión del Aprendizaje.
- **MOOC.** – *Massive Open Online Courses*: Cursos en Línea Masivos y Abiertos.
- **MIT.** – Massachusetts Institute of Technology.
- **SoLAR.** – Sociedad de Investigación de Learning Analytics.
- **EDA.** – Exploratory Data Analysis: Análisis Exploratorio de Datos.
- **CMS.** – Sistema de manejo de contenido
- **OLX.** – XML abierto para el aprendizaje.
- **SRL.** – Self-Regulating Learning: Aprendizaje Auto-Regulado.
- **PCA.** – Principal Component Analysis: Análisis de componente principales.
- **t-SNE.** – t-stochastic neighbor embedding.
- **SDK.** – Software Development Kit: Kit de desarrollo de software.
- **PM.** – Process Mining: Minería de Procesos.
- **ML.** – Machine Learning: Aprendizaje de Máquina.
- **EFLA.** – Evaluation Framework for Learning Analytics: Marco de evaluación para análisis de aprendizaje.
- **LATUX.** – Learning Awareness Tool – User eXperience: Herramienta de concienciación sobre el aprendizaje de la experiencia del usuario.
- **CEDIA.** – Corporación Ecuatoriana para el Desarrollo de la Investigación y Academia.
- **EDA.** – Exploratory Data Analysis: Analisis exploratorio de los datos.
- **GCM.** – Group Concept Mapping: Mapeo de grupo de concepto.
- **TAM.** – Technology Acceptance Model: Modelo de aceptación tecnológica.
- **XLEA.** – XBlock for LEarning Analytics.

Agradecimiento

Agradezco a mi familia por darme su apoyo incondicional en todos los años de la carrera por siempre brindarme sus palabras de apoyo cuando lo más necesitaba y ayudarme a sacar esa fuerza para superar mis límites y nunca rendirme ante ninguna adversidad.

A Dios por guiarme en este gran camino dándome esa fuerza y voluntad para seguir adelante a lo largo de mis años de estudio, brindándome esa paz y tranquilidad que necesitaba para tomar las mejores decisiones en los momentos más críticos de mi vida.

A todos los docentes de la universidad que contribuyeron impartiendo sus conocimientos y experiencia profesional que ayudaron en mi formación a despertar esa chispa de curiosidad por los nuevos campos de estudio relacionado a los datos. De manera especial agradezco al ingeniero Jorge Maldonado por el tiempo dedicado e impartir sus conocimiento y guía para la realización de nuestro trabajo de titulación.

A todos los compañeros que pude conocer en la universidad con los que pude compartir grandes momentos de alegría y tristeza haciendo que esta etapa de sea una de las mejores etapas de mi vida.

Jonnathan Campoberde A.

Agradecimiento

Primeramente, quisiera agradecer a las dos mujeres más importantes de mi vida, a mi mamá Miriam y mi abuela que de cariño le llamo Ita, que sin sus esfuerzos y sacrificios yo no estaría en el lugar que me encuentro ahora.

Agradezco a mi familia y amigos cercanos por sus consejos, cariño, y apoyo incondicional durante todos estos años. A mi tío Edgar por brindarme palabras de aliento cuando más las necesitaba.

Finalmente, quiero agradecer a mi tutor de tesis Ing. Jorge Maldonado, ya que sin su ayuda y mentoría este trabajo de titulación no hubiese sido posible. De igual manera, a mi amigo Jonnathan Campoberde por acompañarme durante todo este proceso y demostrarme que para los sueños uno nunca se rinde y debe superar sus límites.

Miguel Á. Macías.

UCUENCA

Dedicatoria

Dedico a este trabajo a Dios quien ha sido el que me ha dado la voluntad y fuerza para superar todas las adversidades que se me han presentado en este camino recorrido y el que me toca por recorrer lleno de desafíos y metas por cumplir.

A toda mi familia quienes han sido un pilar fundamental durante todo este proceso en especial a mi mamá y papá que han estado ayudándome en las diversas adversidades que se me han presentado a lo largo de mi vida universitaria.

A mis amigos que siempre han estado en los buenos y malos momentos, especialmente cuando tenían esa paciencia para compartir sus conocimientos en temas que no entendía o no dominaba.

A mi hermano Steven para que encuentre en mí un ejemplo a seguir y entienda que si bien se nos presentan adversidades en cualquier etapa de la vida con trabajo duro y constancia nada es imposible de cumplir.

Jonnathan Campoberde A.

Dedicatoria

Dedico este trabajo de titulación a mi mamá, ya que sé que mis logros también los consideras como tuyos y quiero que siempre estés orgullosa de mí.

También dedico este trabajo a mis hermanos Steven y Kevin, con el objetivo de motivarlos y celebrar este logro juntos.

Finalmente, a mi abuela Ita, ya que ella siempre ha puesto el bienestar de sus nietos antes que el de ella misma, y es por eso que hoy estoy aquí.

Miguel Á. Macías.

Capítulo 1: Introducción

1.1. Panorama general y justificación

En la actualidad los usuarios al interactuar con los sistemas de información generan una gran cantidad de datos, dejando como resultado una huella digital en estos (Ben Kei, 2016). En el sector educativo, en los últimos años, las plataformas de e-learning se han convertido en una fuente rica de datos, producto de las interacciones de los estudiantes con los recursos de los cursos. Un ejemplo de esto, son los Cursos Masivos Abiertos y en Línea, conocidos en inglés como MOOC. Estos cursos atraen a una gran cantidad de estudiantes de todo el mundo (Shah, 2020a). Sus altos niveles de aceptación a nivel mundial (antes y durante la pandemia) se ve reflejado en los más de 110 millones de estudiantes que se han registrado en alguno de los 13.5 mil MOOC según el observatorio Class Central (Dhawal Shah, 2020). Este observatorio destaca a las plataformas Coursera y edX como las plataformas más populares y que atrae la mayor cantidad de estudiantes.

Con el propósito de democratizar el acceso a los MOOC, edX liberó el código fuente de su plataforma denominado Open edX, que cuenta con funcionalidad similar a la de la plataforma edX para alojar y desplegar MOOC. Sin embargo, a diferencia de edX, la versión abierta Open edX no cuenta con el módulo de analítica de aprendizaje denominado insights. Este módulo permite analizar los datos que son resultados de las interacciones de los estudiantes en el MOOC y utilizar esa información para extraer conocimiento (Torre et al., 2020). Esto ha dado como resultado, que estudiantes no puedan conocer cuáles han sido sus interacciones durante el proceso de aprendizaje y los profesores encuentren limitaciones al momento de hacer un seguimiento y monitoreo de las actividades de los estudiantes. Como resultado, la retroalimentación es escasa y el apoyo a los estudiantes es casi nulo (Antonaci et al., 2018).

Las Analíticas de Aprendizaje (LA – Learning Analytics en inglés) permiten mejorar los procesos de enseñanza y aprendizaje por medio de la recolección, análisis y entendimiento de los datos sobre los contextos en los que se producen con el fin de mejorarlos. Para Trafford y Shirota (Trafford & Shirota, 2011), las LA permiten escalar las habilidades de los docentes para dar seguimiento a los estudiantes y retroalimentarlos de forma adecuada. Open edX en un intento por incorporar visualizaciones desarrolló la herramienta externa Insights (edX, 2018). Esta herramienta es un dashboard de visualizaciones que se centra solo en los administradores de un curso. Las visualizaciones que se han implementado en Insights han sido catalogadas como muy básicas haciendo que existan proyectos que traten de cubrir esta necesidad (Torre et al., 2020). Un ejemplo de lo anterior es ANALIZE. Este proyecto, incorpora visualizaciones con el objetivo de conocer las interacciones de cada estudiante dentro de la plataforma. Otro ejemplo es

OXALIC, este proyecto provee visualizaciones que se presentan solo a los docentes, pero únicamente información general de los estudiantes y su progreso en el curso (Khalil & Belokrys, 2020). Los proyectos antes mencionados aportan visualizaciones de manera general a los docentes sin que los estudiantes conozcan sobre su proceso de aprendizaje. Esto demuestra que existe aún la necesidad de herramientas que presenten de forma gráfica y visual las acciones y el comportamiento de los estudiantes en este tipo de cursos. Es por esto, que las visualizaciones y el análisis exploratorio de las interacciones de los estudiantes juegan un papel importante al momento de conocer el comportamiento y las estrategias de aprendizaje que emplean dentro de un MOOC (Purwoningsih et al., 2019a).

En este trabajo de titulación se propone desarrollar un componente que implemente visualizaciones que apoyen al proceso de aprendizaje de los estudiantes y al seguimiento y monitoreo como ayuda a los profesores. Para esto se desarrollará un XBlock para la plataforma Open edX que incorporará un conjunto de visualizaciones que den cuenta del proceso de aprendizaje de los estudiantes en un MOOC. Para esto, como parte del proceso, se realizará un estudio exploratorio del comportamiento de los estudiantes para entender qué tipo de visualizaciones son posibles de diseñar a partir de las distintas fuentes de datos. Lo anterior facilitará que los MOOC puedan ser utilizados como material digital educativo de acompañamiento en un curso tradicional o que puedan ser utilizados como materiales autónomos promoviendo procesos de autorregulación del aprendizaje en los estudiantes.

1.2. Estado del Arte

1.2.1. Estudio del comportamiento de los estudiantes en un MOOC

En los últimos años se ha dado un crecimiento en cuanto a la cantidad y uso de los sistemas de gestión de aprendizaje (Learning Management System - LMS por sus siglas en inglés) y su adopción en la educación superior (edutechnica, 2018), más aún con el surgimiento de la pandemia COVID-19. LMS es un término que es usado por abarcar varios sistemas que proveen servicios de educación virtual para estudiantes, instructores y administradores. Los LMS permiten registrar usuarios, mostrar los cursos en su catálogo y además registran lo que los estudiantes realizan al interactuar con los recursos del curso (Aldiab et al., 2019). Lo anterior abre una oportunidad para estudiar y entender los comportamientos de los estudiantes en la plataforma y las secuencias de aprendizaje que siguen (Seaton et al., 2014).

El estudio del comportamiento de los estudiantes dentro de los MOOC se ha abordado desde varias perspectivas, por ejemplo desde el diseño del MOOC (Watson et al., 2016), los subgrupos de personas y su compromiso con los recursos educativos del curso (Kizilcec et al., 2013), la motivación e intereses que tienen los estudiantes (Daphne Koller et al., 2013), entre otras. Sin embargo, estos enfoques no consideran cuales son las interacciones de los estudiantes dentro de la plataforma, esto con el objetivo de evidenciar las secuencias de aprendizaje y que están relacionadas con el éxito académico o terminación del MOOC.

El análisis del comportamiento de los estudiantes desde una perspectiva del constructivismo personal es abordado por varios autores (van den Beemt et al., 2018b). Por ejemplo, en un estudio realizado por Van den Beemt se evidencia una búsqueda por encontrar una forma de entender el comportamiento de los estudiantes dentro de un MOOC. Este estudio como respuesta, plantea el estudio del comportamiento de los estudiantes por medio de la minería de procesos (PM – Process Mining en inglés). Como resultado de este estudio, se mostró este enfoque como fructífero y capaz de revelar el comportamiento de los estudiantes durante su proceso de aprendizaje dentro del MOOC. Este enfoque basado en PM permite conocer a lo largo del tiempo como se ordenan las actividades de los estudiantes y descubrir los patrones que están siguiendo.

Otro estudio importante es el de Barba y colegas. Ellos estudiaron como los estudiantes organizan su tiempo en términos de sesiones o bloques de tiempo a lo largo de un MOOC (Paula G. de Barba et al., 2020). Este estudio analiza los eventos que son generados por los estudiantes, los mismos que se encuentran almacenados en los logs y son usados para determinar las sesiones a nivel de semanas. Las sesiones son analizadas en cuanto a su duración y a la actividad que están realizando. En el primer caso estas son catalogadas como pequeñas, medianas y largas; en el segundo caso estas son divididas dependiendo el porcentaje de tiempo empleado para realizar una actividad específica. El analizar el comportamiento de los estudiantes empleando las sesiones de la manera antes mencionadas evidenció resultados favorables, además se descubrió que los indicadores de sesiones son indicadores significativos en las habilidades del comportamiento autorregulado (Self-Regulated Learning – SRL por sus siglas en inglés). Además, se pudo identificar la conducta que tienen los estudiantes exitosos y estudiantes con altos niveles de SRL, esto como un punto de partida para realizar intervenciones en estudiantes con bajo rendimiento.

1.2.2. Dashboard de Visualizaciones en los MOOC

Las analíticas de aprendizaje ayudan a los estudiantes a entender las actividades que han realizado dentro de un MOOC (Khalil et al., 2015). Estas actividades recopiladas son analizadas y utilizadas para prevenir los riesgos de deserción en el aprendizaje y mejorar el rendimiento de los estudiantes en el curso (Cross et al., 2019; Purwoningsih et al., 2019a). Los datos que muestren la información de los estudiantes comúnmente se muestran a través de visualizaciones agrupadas en un dashboard (Charleer et al., 2016). Un ejemplo de esto es el trabajo de Kovanovic y colegas. Ellos mostraron (Kovanović et al., 2017; Rohloff et al., 2019) que el dashboard mostrado a los estudiantes incrementó la motivación por planificar mejor su aprendizaje, apoyándolos e involucrándolos a aplicar el aprendizaje autorregulado.

Otros ejemplos de implementaciones de dashboard en las plataformas que proveen MOOC son: 1) ALAS-KA, este es un dashboard que presenta visualizaciones del progreso del estudiante dentro de la plataforma, la distribución del tiempo en ejercicios y videos (Ruipérez-Valiente et al., 2015); 2) edX cuenta con ELAT una herramienta que presenta visualizaciones sobre las rutas de aprendizaje, participación en el foro y secuencias en los videos (Torre et al., 2020); 3) VisMOOC que ofrece visualizaciones de los clics y acciones de los usuarios entre cada video e información estadística de los estudiantes (Conglei Shi et al., 2015); 4) moocRP que ofrece visualizaciones del flujo y estructura del curso y estadísticas básicas de los estudiantes (Pardos et al., 2016).

Open edX en un intento por incorporar visualizaciones dotadas de analíticas de datos desarrolló la herramienta externa Insights (edX, 2018). Edx Insights proporciona datos sobre la actividad y el rendimiento de los alumnos en un curso. Entre las visualizaciones que proporciona se muestran métricas de las inscripciones de tipo geográfica, demográfica; participación en el contenido del curso en semanas de manera general; cantidad de veces que se reproducción un video de una sección; cantidad de alumnos que han enviado una respuesta correcta e incorrecta. Estas visualizaciones las pueden acceder solo los administradores de un curso, además estas han sido catalogadas como muy básicas haciendo que existan proyectos que traten de cubrir las necesidades que insights no cubre (Torre et al., 2020).

En relación con Open edX, uno de los primeros desarrollos que implementa dashboards de analítica es ANALYSE (Add-on of learNing AnaLYtics Support for open Edx – ANALYSE por sus siglas en inglés) (Ruipérez-Valiente, Muñoz-Merino, Pijeira Díaz, Ruiz, et al., 2017). ANALYSE se encarga de procesar los

UCUENCA

datos generados por los estudiantes y provee visualizaciones para estudiantes y profesores por medio de su dashboard. Las visualizaciones que están incorporadas en su dashboard muestran analíticas resumen del curso, notas de los estudiantes en cada actividad, tiempo que se toman los estudiantes para resolver un problema, tiempo empleado para observar los videos, los eventos que el usuario han generado mientras observan los videos, los accesos de los estudiantes a la plataforma, el tiempo empleado en cada capítulo. Estas visualizaciones han sido evaluadas en los aspectos de usabilidad y efectividad, mostrando que han tenido buena aceptación por parte de los estudiantes y profesores, pero actualmente esta se encuentra obsoleta, dado el versionamiento de open edX.

Finalmente, OXALIC es una herramienta más reciente que provee analíticas de aprendizaje para la plataforma open edX (Khalil & Belokry, 2020). Esta herramienta se encarga de realizar un procesamiento a los logs de eventos que son obtenidos de la plataforma para mostrar diferentes visualizaciones, entre las cuales se tiene estadísticas generales del curso, foros con los que interactúa el estudiante, número de personas que interactúan con los videos, eventos que se generan al ver los videos, el tiempo que el estudiante se encuentra en la plataforma y un mapa que representa las actividades que los estudiantes han realizado en la plataforma. En la **Tabla 1** se muestra un resumen con el detalle de los trabajos antes mencionados.

Tabla 1. Vista general de las herramientas de visualización que existen para Open edX.

Estado actual	ANALYSE	OXALIC	INSIGHTS
Es un componente XBlock	No	No	No
Se encuentra funcionando	Obsoleto	Privado	Sólo para algunas versiones de Open edX
Estudiantes pueden acceder a las visualizaciones	No a todas	No	No
Vista general del curso	Sí	Sí	Sí
Análisis del curso por día/semana	No	Sí	Sí

Análisis del comportamiento de los estudiantes	No	Sí	No
--	----	----	----

Los trabajos antes mencionados muestran las visualizaciones que han sido consideradas por ellos como las más idóneas, pero estas no garantizan que las visualizaciones mostradas en su dashboard sean las mejores debido a que no se ha considerado las variables que influyen en el rendimiento académico y que están relacionadas con las secuencias de aprendizaje de los estudiantes (Maldonado-Mahauad, Pérez-Sanagustín, Moreno-Marcos, et al., 2018). Por otro lado, ninguno de los proyectos analizados ha sido desarrollados como un componente XBlock, que permita extender la funcionalidad de la plataforma open edX e independizarlo de la versión de la plataforma para MOOC. Por estos motivos, en este trabajo de titulación se presenta el desarrollo de un componente XBlock para la plataforma Open edX.

1.3. Objetivos del Trabajo

1.3.1. Objetivo general

Desarrollar un componente XBlock para la plataforma Open edX que incorpore visualizaciones sobre el comportamiento de los estudiantes para brindar información sobre el proceso de aprendizaje en un curso MOOC.

1.3.2. Objetivos específicos

- Desarrollar un análisis exploratorio del comportamiento de los estudiantes para determinar variables y secuencias de aprendizaje realizadas por los estudiantes en un MOOC.
- Diseñar un dashboard de visualizaciones para profesores y estudiantes que den cuenta del comportamiento de los estudiantes.
- Implementar un XBlock para Open edX para visualizar el comportamiento de los estudiantes en un MOOC.
- Evaluar la usabilidad de las visualizaciones desarrolladas en el XBlock.

1.4. Metodología

La metodología empleada para este trabajo de titulación consta de cinco etapas (ver **Figura 1**) adaptadas de la metodología LATUX (Learning Awareness Tool – User eXperience) (Martinez-Maldonado et al., 2016). Esta metodología es útil en los contextos donde el flujo de trabajo para el diseño e implementación de herramientas mejoren la entrega de información en entornos de aprendizaje

UCUENCA

virtual. Las cinco etapas se agrupan bajo 2 enfoques: 1) el primer enfoque se centra en la identificación del problema y sus indicadores, aquí se debe realizar un análisis exploratorio e identificación de variables que ayuden a mejorar el éxito estudiantil en plataformas de aprendizaje en línea; y 2) el segundo enfoque aborda las etapas de planificación, diseño, implementación y evaluación. La evaluación de las visualizaciones incorporadas en el XBlock será realizada por estudiantes y profesores de un curso MOOC de la Universidad de Cuenca. Los participantes tendrán acceso completo a todas las visualizaciones que serán incorporadas, tomando en cuenta su rol dentro de la plataforma se proporcionará el cuestionario respectivo. Para medir la calidad del componente XBlock como herramienta para el soporte de LA, se hará uso del cuestionario Evaluation Framework for Learning Analytics (EFLA) (Scheffel et al., 2017), realizando la evaluación del sistema como un todo y también las visualizaciones de manera individual. Adicionalmente, se aplicará la evaluación User Experience Questionary (UEQ) (Laugwitz et al., 2008), para medir la experiencia de usuario de los participantes al usar nuestro componente de LA.

UCUENCA

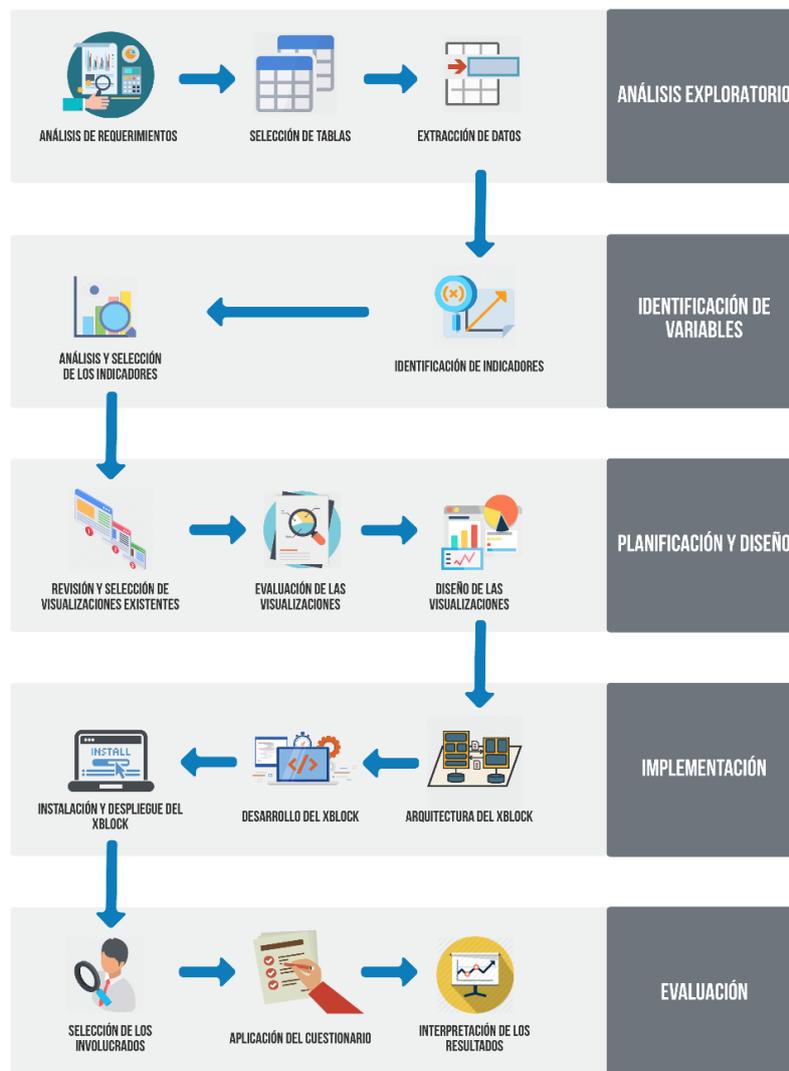


Figura 1. Pasos de la metodología adaptada LATUX a emplear para el desarrollo del trabajo de titulación.

A continuación, se detallan los pasos de cada etapa de la metodología:

- **Análisis exploratorio:** Se requiere realizar un análisis de los requerimientos del diseño de las visualizaciones, ya que es fundamental centrarse en lo que los estudiantes y maestros necesitan para mejorar el proceso de aprendizaje. Además, se deberá estudiar la estructura de un curso de Open edX con el propósito de seleccionar las tablas que brinden información relevante sobre el comportamiento de los estudiantes.

UCUENCA

- **Identificación de variables:** Para identificar los indicadores que permitan mejorar el éxito estudiantil se realizó una descomposición de los requerimientos obtenidos previamente. Además, se referencian los indicadores obtenidos en otros trabajos con el propósito de seleccionar las variables más relevantes para este trabajo de titulación.
- **Planificación y Diseño:** Luego de haber obtenido los indicadores de éxito estudiantil se realizó una revisión exhaustiva sobre visualizaciones ya existentes que aporten analíticas de aprendizaje, con el propósito de seleccionar las mejores y adaptarlas a nuestro trabajo de titulación. Posteriormente, se evaluaron estos diseños para asegurarse que cumplan con los objetivos establecidos.
- **Implementación:** Una vez han sido diseñadas las visualizaciones que se van a incorporar, se procede a desarrollar el componente de XBlock que será integrado dentro de un curso de Open edX. Para esto, primero se deberá estudiar la arquitectura de un XBlock y de la plataforma.
- **Evaluación:** La evaluación de las visualizaciones incorporadas en el XBlock será realizada por estudiantes y profesores de un curso MOOC de la Universidad de Cuenca. Los participantes tendrán acceso completo a todas las visualizaciones, tomando en cuenta su rol dentro de la plataforma se proporcionará el cuestionario respectivo

1.5. Estructura del trabajo de titulación

Este trabajo de titulación se estructura de la siguiente manera:

El capítulo 2, introduce al lector a la definición de MOOC, sus características, funcionalidades, y los patrones que se pueden obtener de los datos que dejan los estudiantes en la plataforma y las herramientas disponibles para tratar estos datos y convertirlos en información. Además, presenta como se desarrolla el proceso de aprendizaje de los estudiantes dentro de los MOOC.

En el capítulo 3, se describe a detalle la arquitectura del sistema de gestión de aprendizaje empleado por Open edX. Se presenta la estructura de los cursos y la gestión de los datos dentro de la plataforma, es decir, base de datos y archivos de rastreo de información de las interacciones de los estudiantes.

El capítulo 4 aborda el análisis exploratorio del comportamiento de los estudiantes dentro de los MOOC. Para este trabajo de titulación se hace especial énfasis en los datos y patrones obtenidos a partir de la plataforma Open edX.

UCUENCA

Además, se presenta la revisión de la literatura con otros trabajos con el propósito de identificar las variables relacionadas con el éxito académico, como así también, visualizaciones que aporten analíticas a estudiantes y profesores. Finalmente, se revisa la estructura del log de eventos dentro de Open edX y la manera en la que se extraen y procesan estos datos.

En el capítulo 5, se planifica el diseño de las visualizaciones a ser implementadas en el componente a desarrollar para Open edX. Se revisan herramientas de LA ya existentes y se seleccionan aquellas visualizaciones que se alineen a los objetivos y requerimientos de este trabajo. Las visualizaciones estarán dotadas de los indicadores seleccionados en el capítulo 4.

En el capítulo 6, se presenta el desarrollo del componente XBlock y la implementación de las visualizaciones. Para esto primero se presenta la arquitectura de un XBlock y se definen las funcionalidades que va a incorporar para las visualizaciones. El capítulo finaliza con la evaluación del componente XBlock y sus resultados.

Finalmente, en el capítulo 7 de este trabajo de titulación, se presentan las conclusiones, recomendaciones y trabajos futuros. Se revisan los objetivos comprometidos, así como también las recomendaciones para seguir trabajando en esta línea de interés y se sugiere futuras mejoras.

Capítulo 2: Proceso de aprendizaje dentro de los MOOC

2.1. Introducción

En esta sección se presenta una descripción de los conceptos empleados en este trabajo de titulación, con el fin de que el lector se familiarice con el área de estudio en el que se desarrolla este trabajo de titulación. Se explica qué son los MOOC, sus inicios y el impacto que han tenido en la última década. Así también, se abordan conceptos acerca del comportamiento de los estudiantes dentro de los MOOC, y las técnicas utilizadas para analizar sus patrones y secuencias de aprendizaje. Finalmente, se exploran las posibles variables e indicadores que permitan mejorar el éxito estudiantil.

2.2. ¿Qué son los MOOC?

El concepto de MOOC hace referencia a sus iniciales en inglés (Massive Open Online Course), también en el idioma español existe una referencia por su acrónimo COMA (Cursos Online Masivos Abiertos), CAEM (Curso Abierto En línea Masivo), o CALGE (Curso Abierto en Línea a Gran Escala), es decir no existe todavía una estandarización y se lo conoce más a nivel mundial como MOOC (Atiaja et al., 2016). Se caracterizan (**Figura 2**) por ser cursos en-línea, no solicitar requisitos formales para registrarse, no tener límites de participación, no realizar ningún cobro monetario y no ganar créditos (Gaebel, 2013).

Dentro de las distintas definiciones, muchas de ellas basándose en su acrónimo, encontramos las siguientes planteadas por diferentes autores como; McAuley, quien define a los MOOC como un fenómeno que integra la conectividad de las redes sociales, el acceso de expertos reconocidos en un campo de estudio (tutores) y una colección de recursos abiertos e libre acceso en línea, donde varios estudiantes (cientos o miles), pueden participar de manera activa, donde ellos se ajustan a los objetivos, conocimientos previos, habilidades e intereses comunes de aprendizaje (McAuley et al., 2010). Otra definición planteada, la realiza por Littlejohn, (*Patterns of Engagement in Connectivist MOOCs - Enlighten: Publications*, n.d.), en la cual argumenta que un MOOC es un curso destinado a la participación interactiva a gran escala (masiva o de muchos usuarios) y al acceso abierto a través de internet, considerando una forma relativamente nueva de aprendizaje.

Se puede apreciar entonces que la definición de un MOOC, se basan básicamente en su acrónimo, pero en muchas ocasiones puede generar controversia. Sin embargo, no hay una definición más allá confirmada o adoptada como estándar, lo que ha llevado a los MOOC a ser confundidos con repositorios de recursos en línea, pero su diferenciador radica en la posibilidad en el formato de curso de los MOOC, el cual fomenta la participación de los estudiantes en

para participar en actividades de aprendizaje que permiten alcanzar los objetivos planteados por el material educativo digital.

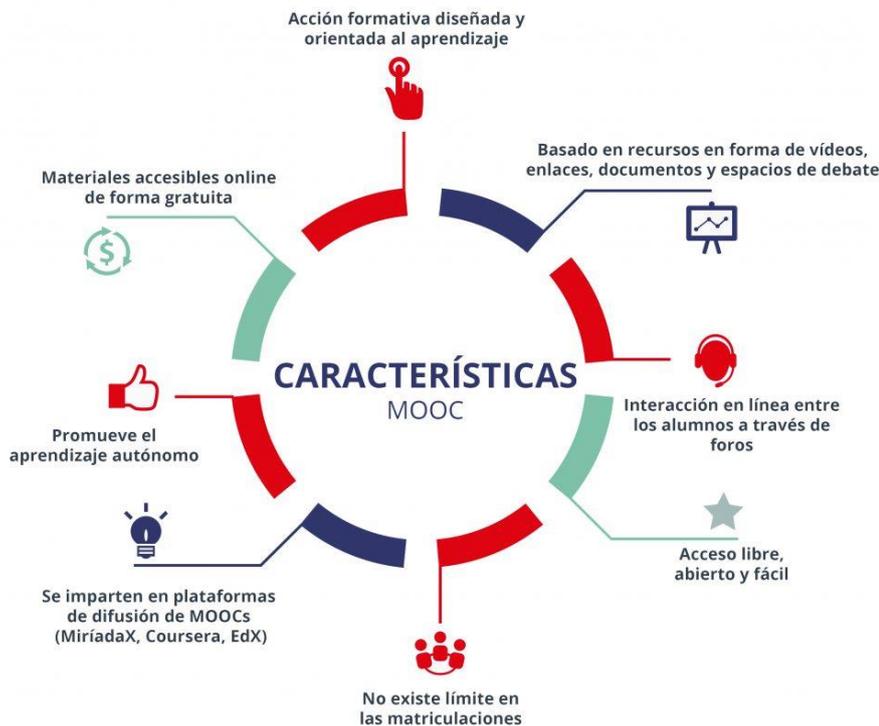


Figura 2. Características de los MOOC.

Nota: Tomado de (UBUCEV, 2019).

Los MOOC nacen con la idea inicial de la educación a distancia, y son producto de la evolución de la misma. Si se considera la evolución de la educación a distancia, donde se entrega el material o recursos (libros, cuadernos, textos guía, etc.), para que el estudiante lo revise en su hogar y sin restricción de horario con entregas periódicas y clases presenciales cada cierto tiempo obligatorias. Al continuar con esta evolución se presenta el primer avance; el audiovisual, donde los maestros grabaron audios y videos con el motivo de mejorar la experiencia de interacción, pero de la misma forma, la entrega del material o recursos era física (VHS, CD, Casete) (Liyanagunawardena et al., 2013)

Los MOOC como término formal, apareció en 2008 y fue propuesto por Dave Cormier (Smith & Eng, 2013), desde entonces estos se han convertido en el catalizador del cambio del modelo tradicional de enseñanza/aprendizaje de las universidades.

UCUENCA

El primer curso MOOC que abrió el camino de esta modalidad de aprendizaje fue denominado “Connectivism and Connective Knowledge - Conectivismo y aprendizaje conectivo”, con una duración de 12 meses y una participación de aproximadamente 2 300 estudiantes de distintos países, quienes se registraron en forma libre. Este curso fue organizado por George Siemens y Stephen Downes en la Universidad de Manitoba (Canadá) en el 2008.

En el 2011, se lanza el curso “Introduction to Artificial Intelligence - Introducción a la Inteligencia Artificial”, que fue organizado por Sebastián Thrun, profesor de la Universidad de Stanford y Peter Norvig, director de investigación de Google, en este curso se registraron 160 000 estudiantes de todo el mundo (Atiaja et al., 2016).

Desde entonces varias universidades en el mundo se han puesto en marcha en lo que respecta a la producción de MOOC (ver **Figura 3**). En el año 2020 se registraron alrededor de la tercera parte de todas las personas que alguna vez se han registrado en un MOOC, convirtiendo al 2020 en “El Segundo año de los MOOC” (Shah, 2020c). Los datos recolectados del observatorio mundial de MOOC, Class Central, muestran que solo en el 2020 se crearon alrededor de 2800 cursos, 360 micro credenciales y 19 títulos en-línea.

En los 10 años de historia, los MOOC se conceptualizan como la evolución del aprendizaje en línea. Además, su popularidad se explica a partir del aval de algunas de las universidades más prestigiosas de Estados Unidos y el mundo, como, por ejemplo; Universidad de Stanford en E.U., Universidad de Edimburgo en Reino Unido, la Pontificia Universidad de Chile en Latinoamérica, la Universidad China de Hong Kong, entre otras las mismas que han acogido y desarrollado con entusiasmo el concepto a través de iniciativas como EdX, Coursera o Udacity. (Torres, 2013).

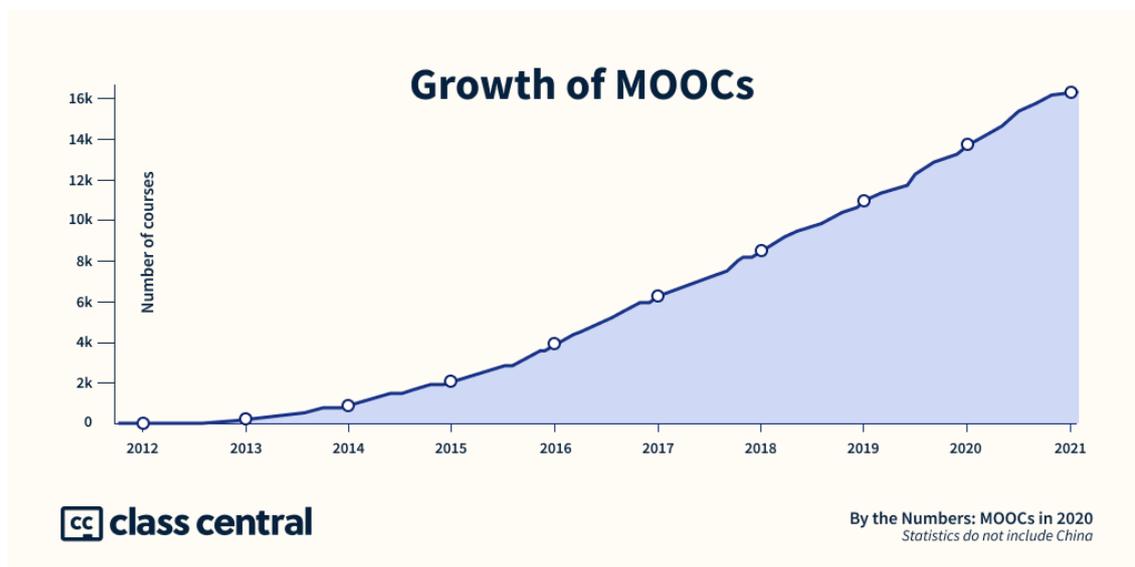


Figura 3. Crecimiento de los MOOC a través de los años.

Nota: Tomado de Shah (2020b).

En el estudio de Loeckx (Loeckx, 2016) asegura que el aprendizaje en línea se ha convertido en un prioridad estratégica. Las universidades han tomado una postura a favor de los cursos en línea, con el fin de mejorar la visibilidad de su institución a través de los MOOC. Debido al crecimiento de los MOOC, han surgido ramificaciones en lo que respecta a la manera de clasificarlos, sin embargo podemos distinguirlos en dos grandes grupos: MOOC conectivistas (cMOOC) y los MOOC tradicionales (xMOOC) (Smith & Eng, 2013). La última es conocida como la nueva generación de MOOC, denominados xMOOC por edX, en ese entonces MOOC de Stanford (Jim, 2012) surgieron en el otoño de 2011 cuando la Universidad de Stanford ofertaba un curso gratuito en línea acerca de Inteligencia Artificial. Estos cursos hacen uso de un diseño “instructivita”, el cual se basa en objetivos de aprendizaje predefinidos por un guía o instructor, donde los estudiantes interactúan de forma limitada entre ellos.

2.3. Definición de patrón de aprendizaje en un MOOC

El entorno de aprendizaje en el que se desenvuelven los estudiantes puede contener varios factores que afecten la manera en la estudian o aprenden (Vermunt, 2005), tales como la motivación del estudiante, hábitos y habilidades de estudio. Esto va de la mano con las actividades de aprendizaje que el estudiante debe realizar dentro de un curso, lo cual puede llegar a determinar la calidad de los resultados de aprendizaje que consigan.

Para llegar a comprender el comportamiento de los estudiantes dentro de los MOOC, es importante conocer acerca de los patrones de aprendizaje (**Figura 4**). Vermunt en su estudio (Vermunt & Donche, 2017), explica que, un patrón de aprendizaje se conceptualiza como un conjunto coherente de actividades de aprendizaje (secuencias) que suelen emplear los alumnos, sus creencias y motivaciones de aprendizaje en un determinado periodo de tiempo.

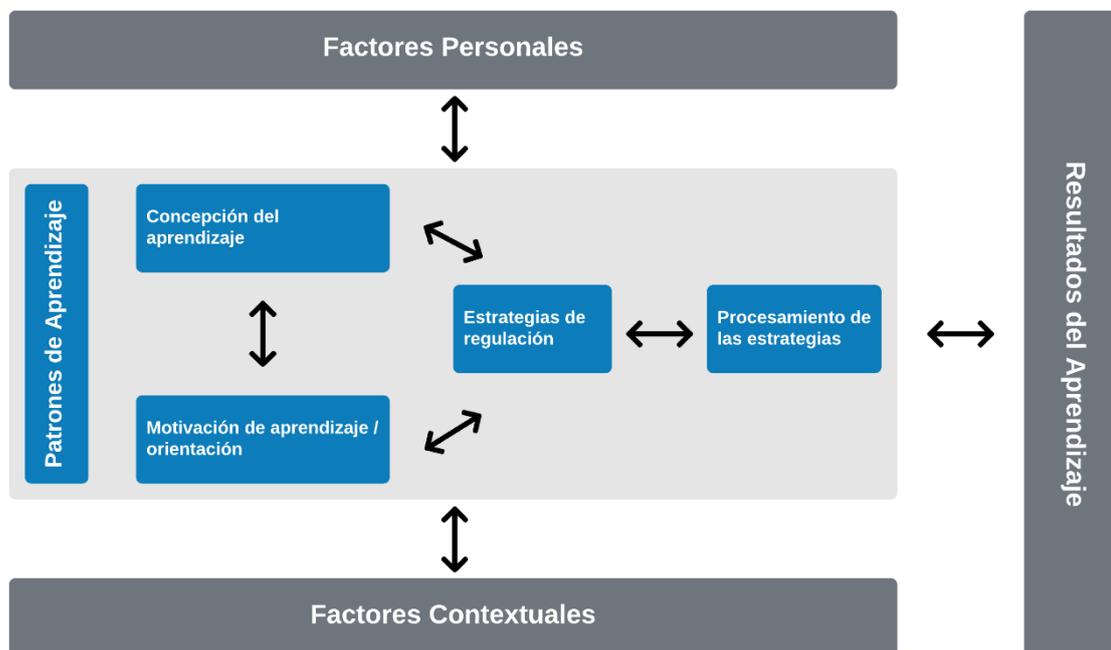


Figura 4. Modelo de los patrones de aprendizaje.

Nota: Adaptado al español de (Vermunt & Donche, 2017).

Estos patrones pueden ser capturados a través del análisis de logs que contengan las interacciones secuenciales del comportamiento de los estudiantes, mediante el uso de técnicas de minería de procesos sobre los datos educativos (Boroujeni & Dillenbourg, 2018) y se puede lograr un apoyo más efectivo y personalizado durante el proceso de aprendizaje dentro de un curso MOOC. Esto representa una ventaja sobre la explotación de las capacidades computacionales de hoy y que podrían revelar patrones ocultos en las actividades de los estudiantes y proporcionar información acerca de sus estrategias de aprendizaje dentro de los SGA.

2.3.1. Minería de Procesos en un contexto educativo

La Minería de Procesos es una de las técnicas que se utilizan para analizar los datos generados en los procesos de aprendizaje (**Figura 5**) utilizando los datos

educativos (Van Der Aalst et al., 2004). Esta técnica, comúnmente empleada en el área empresarial, se ha ido adaptando en diferentes áreas (p.e., medicina, educación, marketing) para extraer los patrones de comportamiento. Para el caso de los MOOC, la MP es una técnica que promete mucho. Por ejemplo, en un trabajo de Mukala, se reporta (Mukala et al., 2015) que los estudiantes exitosos tienen más probabilidades de estudiar secuencialmente que los estudiantes menos exitosos.

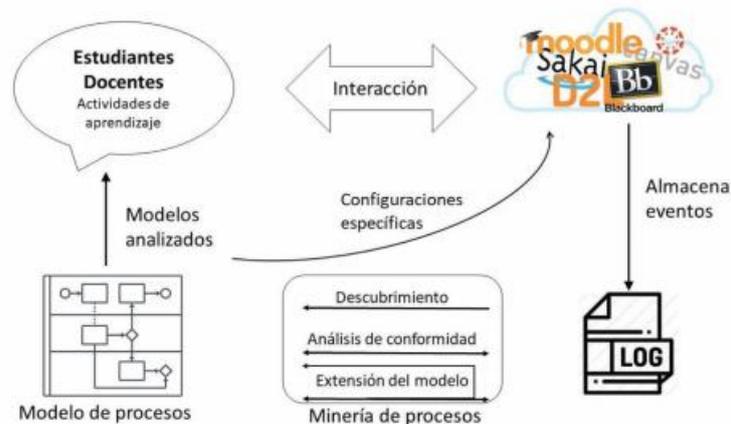


Figura 5. Elementos de la Minería de Procesos en un contexto educativo.

Nota: Tomado de (Fernández Buenaño & Luján-Mora, 2019).

En otro trabajo presentado por Van den Beemt (van den Beemt et al., 2018b), utilizando MP muestran la forma en la que estudiantes aprobados y reprobados distribuyen sus actividades a lo largo de las semanas de un curso. Los grupos de estudiantes que pudieron identificar revelan que, los estudiantes exitosos exhiben un comportamiento de aprendizaje más estable. Es así que, determinar los patrones de aprendizaje a partir de secuencias de actividades en los MOOC utilizando MP se está volviendo común para comprender mejor el proceso de aprendizaje (Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, et al., 2018).

2.3.2. Descubrimiento de Patrones mediante el Aprendizaje de Máquina

Otra de técnica para el descubrimiento de patrones que se ha vuelto popular en los últimos años es el Aprendizaje de Máquina (ML – por sus siglas en inglés). El ML es una técnica eficaz que se puede aplicar junto con las LA con el propósito de descubrir patrones ocultos en la interacción de los estudiantes con los MOOC (Al-Shabandar et al., 2017). La aplicación de técnicas de ML tales como K-means y PCA han demostrado ser eficientes al momento de entender los patrones de

comportamiento de los estudiantes (Kloft et al., 2014; Purwoningsih et al., 2019a).

El uso de PM y ML para analizar el comportamiento de los estudiantes dentro de los MOOC trae consigo beneficios, especialmente para revelar, mapear y monitorear procesos orientados al aprendizaje. Tal como se ha explicado en esta sección, estas técnicas resultan ser útiles como herramientas analíticas en el contexto del aprendizaje en línea.

2.4. Variables relacionadas con el éxito en los MOOC

En muchas de las teorías que se tiene sobre el éxito del aprendizaje de un individuo implican que el conocimiento previo es clave (Kennedy et al., 2015). En el estudio de (Council, 2000), afirma que los estudiantes no solo tienen conocimientos previos sino también cuentan con habilidades preexistentes que aportan a la manera en la que aprenden. Las personas que participan en un MOOC por lo general deben estar preparadas para gestionar su propio aprendizaje (p.e., autorregulados) en base a sus habilidades y conocimientos previos, pues estos cursos cuentan por lo general con una estructura curricular abierta por lo que pueden completar tareas en cualquier orden (Kennedy, 2015).

El aprendizaje en los MOOC es uno de los temas que ha creado debate, dividiendo a las personas en dos formas de pensamiento. La primera, donde las personas defienden que esta nueva forma de enseñanza podría dejar obsoleta a las universidades; y la segunda, están las personas que piensan que no funciona en muchas disciplinas académicas, debido a que los estudiantes están más acostumbrados al ambiente que se tiene en las aulas (Audrey Watters, 2013). Estas formas de pensamiento puede estar relacionadas en la forma como se realiza la medición del éxito en los entornos de aprendizaje en línea (Williams et al., 2018).

El éxito de los estudiantes en los MOOC tiene diferentes métricas para medir y predecir los resultados, tomados desde diferentes perspectivas (Gardner & Brooks, 2018). Las métricas que más comúnmente están relacionadas al éxito de los estudiantes se encuentran: a) el abandono o deserción, cuando un estudiante no interactúa con el curso hasta que este finaliza (Nagrecha et al., 2017); b) detener el aprendizaje, relacionado a los estudiantes que dejan de participar en el curso antes de que este finalice, comúnmente asociado con la deserción pero con la diferencia de que no se puede observar la intención de un estudiante de abandonar (Taylor et al., 2014); c) la certificación, cuando los estudiantes logran tener un mínimo de puntos requeridos en cada tarea para aprobar y obtienen un certificado de finalización (DeBoer et al., 2014); d) nota del examen final, la calificación que los estudiantes obtienen de la prueba

acumulativa al final del curso Taylor (2014); e) fallo o aprobación, los estudiantes aprueban el curso cuando superan una nota mínima requerida; compromiso de los estudiantes relacionado directamente al comportamiento que los estudiantes muestran en el progreso, participación y finalización de tareas (Bote-Lorenzo & Gómez-Sánchez, 2017).

Las variables que comúnmente son asociadas al éxito de los estudiantes se muestran en los siguientes estudios. En el estudio de (Samuelsen & Khalil, 2020) detalla que el tiempo que los estudiantes se dedican a los videos, la edad de la persona, el sexo, la educación y el esfuerzo son variables asociadas al éxito de los estudiantes en los MOOC. En Kennedy (2015) de un estudio de 6.635 participantes solo pasaron 774 personas, indicando que la actividad general o el compromiso que se tiene en el curso y el conocimiento necesario previo son variables asociadas con el éxito. Además, se evidenció que el número de días que un estudiante se encuentra activo en el curso determina significativamente si este mismo tendrá éxito. Nagrecha (2017) analiza que las variables que comúnmente se eligen son clics realizados, notas de las tareas asignadas, información demográfica y las interacciones para la predicción del abandono de un estudiante.

Conclusiones

En este capítulo se explica el concepto de MOOC, sus inicios alrededor del año 2008 y el impacto que han tenido en la última década en universidades de todo el mundo. Posteriormente, se abordan aspectos sobre el comportamiento de los estudiantes dentro de los MOOC. Este comportamiento da lugar a patrones que pueden ser capturados a través del análisis de logs que los estudiantes generan al interactuar con las actividades de un MOOC y que se convierten en patrones de aprendizaje. En varios estudios una de las técnicas utilizadas para analizar estos datos es la Minería de Procesos, la cual revela información acerca de las secuencias y estrategias de aprendizaje de los estudiantes, por ejemplo, grupos de estudiantes exitosos que exhiben un comportamiento de aprendizaje más estable.

Es a partir de esto, que los investigadores se han visto en la necesidad de buscar formas de predecir el rendimiento de los estudiantes, y encontrar la manera de identificar variables que estén relacionadas al éxito académico de los estudiantes en los MOOC. Las variables que comúnmente se asocian al éxito académico están relacionadas a la cantidad de actividades que interactúan los estudiantes dentro del curso, por ejemplo, lecturas, videos y exámenes.

Capítulo 3: Sistema de Gestión de Aprendizaje Open edX

3.1. Introducción

En este capítulo, se introduce al lector, la arquitectura de la plataforma Open edX, sobre la cual se basa este trabajo de titulación. Se presentan sus características principales, la manera en la que funciona y la estructura de los cursos que despliega.

Finalmente, se detalla la manera en la que Open edX registra las interacciones de sus estudiantes dentro de la plataforma, la cual se almacena en forma de logs y también la base de datos que se maneja.

3.2. Arquitectura de Open edX

Open edX está diseñado para ser una plataforma basada en la web que permite la creación, entrega y el análisis de los cursos en línea. El software usado en Open edX es el mismo que se usa en edx.org, una de las plataformas con mayor acogida de estudiantes a nivel mundial y otros sitios de educación en línea (Shah, 2020b). En esta sección, se explica brevemente cómo está conformada su arquitectura desde una perspectiva de alto nivel sin llegar a entrar en muchos detalles debido a su complejidad.

Los componentes que conforman la plataforma son muy diversos y de diferente tipo, tal como se lo muestra en la **Figura 6**. En esta figura se observa una diversa cantidad de tecnologías empleadas y todas estas se comunican por medio de APIs siempre que se pueda realizar de esta manera; Entre toda esta variedad de tecnologías mostradas destaca el uso del Framework Django, el cual usa el lenguaje de programación Python para su uso, esto es importante debido a que contienen las aplicaciones LMS y Studio, siendo este último el sistema de manejo de contenido denominado CMS por sus siglas en inglés. Estas dos piezas centrales de la plataforma son sostenidas por una colección de servicios de tipo web autónomos denominados como aplicaciones implementadas de forma independiente (IDA), esta división ha permitido manejar la complejidad del código base con el que cuenta la plataforma (Open edX, 2022b).

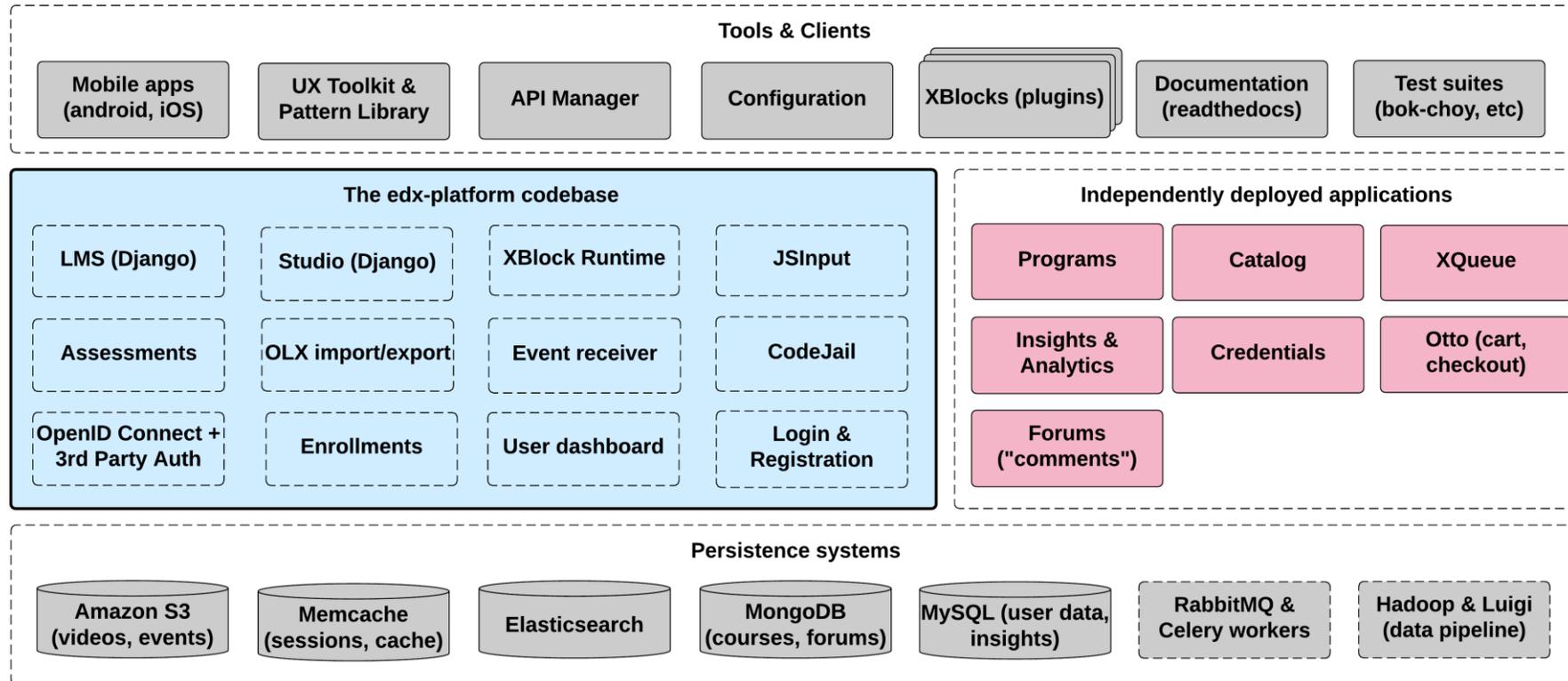


Figura 6. Breve descripción de la arquitectura de Open edX.

Nota: Tomado de (Open edX, 2022b)

A continuación, se da una breve descripción de los componentes claves:

- **LMS:** La parte visible de la plataforma usada por los estudiantes que están inscritos en un curso, además proporciona un panel de instructor. Este es caracterizado por usar las bases de datos de MongoDB para almacenar los cursos, con videos de YouTube o Amazon S3 y MySQL para la información de los estudiantes.
- **Studio:** Es el entorno encargado de la creación de los cursos, este entorno es usado también para actualizar los cursos y sobrescribe la base de datos de los cursos en MongoDB.
- **Discussions:** Este es un IDA encargado de las discusiones con el propósito de brindar al estudiante una mejor experiencia con notificaciones de los temas de su interés.
- **Aplicación Móvil:** Open edX cuenta también con una aplicación móvil tanto para Android como para IOS que es usada por los estudiantes para visualizar los cursos.
- **Analytics:** Encargado de analizar los eventos que genera el estudiante y que son capturados por un archivo JSON, para luego ser procesados en Hadoop y guardados en MySQL. Estos resultados están disponibles para Insights por medio de una API REST.
- **Search:** Búsqueda de información en múltiples contextos incluidos la búsqueda de cursos y comentarios por medio del motor de búsqueda de textos Elasticsearch.

3.3. Estructura de los cursos

Los cursos de open edX están compuesto por unidades llamadas XBlocks, estos XBlocks permiten que los tutores y tecnólogos puedan ampliar el conjunto de componentes que tienen para ofrecer los cursos. Un XBlock es una arquitectura de componentes diseñada para facilitar la creación de nuevos instrumentos de enseñanza dentro de la plataforma, inspirada en las aplicaciones web donde se tienen muchos componentes para ser mostrados en una sola página web. Además de los XBlocks los cursos usan OLX (XML abierto para el aprendizaje) para poder importar y exportar la estructura de los cursos. (Open edX, 2022b)

La estructura de los cursos para la plataforma open edX está definida dentro de edX-Insider. EdX-Insider es un módulo pequeño y un componente principal dentro de la plataforma caracterizado por que la fuente de los cursos tiene un formato XML y los archivos de video, entre otros archivos estáticos son descargados para ser visualizados en la plataforma (edX, 2021). El archivo XML usado para definir la estructura de un curso se llama course.xml. Este archivo se

UCUENCA

caracteriza por tener como elemento principal la etiqueta **course**, dicha etiqueta agrupa la información de la lista de los módulos o XBlocks usados dentro del curso, nombre del curso, la ruta de la URL de acceso, entre otra información respecto al curso; la etiqueta **chapter** es la siguiente en la jerarquía y es la encargada de guardar información respecto al nombre del capítulo o subsección; descendiendo en la jerarquía tenemos a la etiqueta **sequential** usada para todos los elementos secundarios de los capítulos, este contiene información relacionada al nombre de la subsección, la fecha de lanzamiento a los estudiantes, valor booleano si esta subsección es calificada o no, el formato de la nota asignada, entre otra información usada para las calificaciones; en el siguiente nivel de esta jerarquía tenemos a la etiqueta **vertical**, la cual es la encargada de definir una unidad dentro de un capítulo del curso y la información que esta contiene es el nombre de la unidad; finalmente en esta jerarquía tenemos los componentes del curso, la plataforma admite una gran variedad de componentes definidos tales como **html**, **video**, **problem** y componentes personalizados (edX, 2021). Toda esta jerarquía antes descrita se encuentra visualizada en la **Figura 7**.

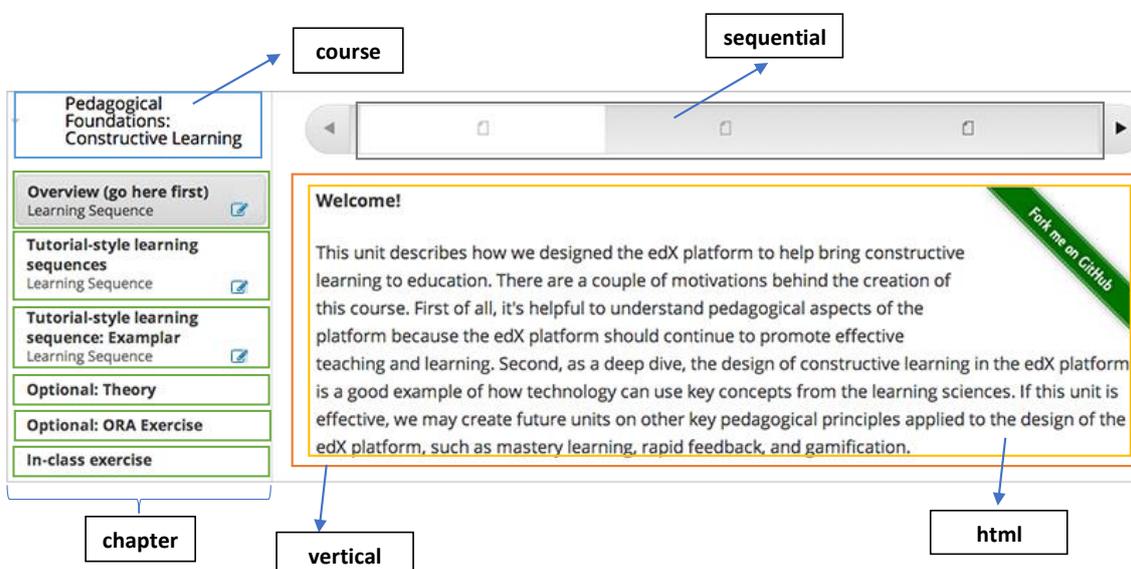


Figura 7. Estructura de un curso en Open edX.

3.4. Gestión de los datos

En esta subsección del capítulo se introduce al lector las tecnologías usadas por la plataforma para la persistencia de los datos, el uso que se da dentro de la misma y por último se describe los archivos de registro de eventos por parte de la plataforma denominado tracking logs.

3.4.1. Base de datos

Open edX almacena los datos de la manera persistente por medio de dos sistemas de gestión de bases de datos encargados del almacenamiento de la información de los estudiantes, cursos, accesos a la plataforma, notas del curso y demás información que sea necesaria para la plataforma (Open edX, 2022b). Estos dos sistemas de gestión son de diferente tipo y cada uno de ellos cumple con un propósito diferente dentro de la plataforma.

MySQL es un sistema de gestión de base de datos relacionales de código abierto y es considerado como la base de datos más popular (Angel Robledano, 2019a). Las características que ha atraído a los desarrolladores a escoger MySQL son su flexibilidad y su fácil uso; alto rendimiento, aunque se estén almacenando una gran cantidad de datos, estos tienen un tiempo de respuesta óptimos; seguro, con un sistema de privilegios de acceso y administración de cuentas de usuarios (Angel Robledano, 2019a). Las características antes mencionadas han hecho que se tome en cuenta esta base de datos para almacenar los datos de los accesos a la plataforma, cursos a los que está inscrito un estudiante, notas de los estudiantes, información demográfica de los estudiantes, entre otra información relacionada a los estudiantes.

La selección de una base de datos diferente para datos más grandes ha llevado a que se seleccione una base de datos de tipo no relacional como es el caso de MongoDB (Open edX, 2022b). MongoDB es una base de datos orientada a documentos, donde los datos no se guardan en tablas, sino en estructuras de datos en forma de documentos denominadas BSON (Binary JSON) con un esquema de tipo dinámico. Entre las características de esta base de datos tenemos que cuenta con un balance de carga para escalar la carga de trabajo, el almacenamiento de los archivos, indexación un concepto similar a las bases de datos relacionales (Angel Robledano, 2019b). En esta base de datos se encarga de almacenar información de la estructura del curso, contenido de los foros de discusión. En la **Figura 8** se muestra cómo se encuentra la arquitectura de open edX y que componentes interactúa con la base de datos.

UCUENCA

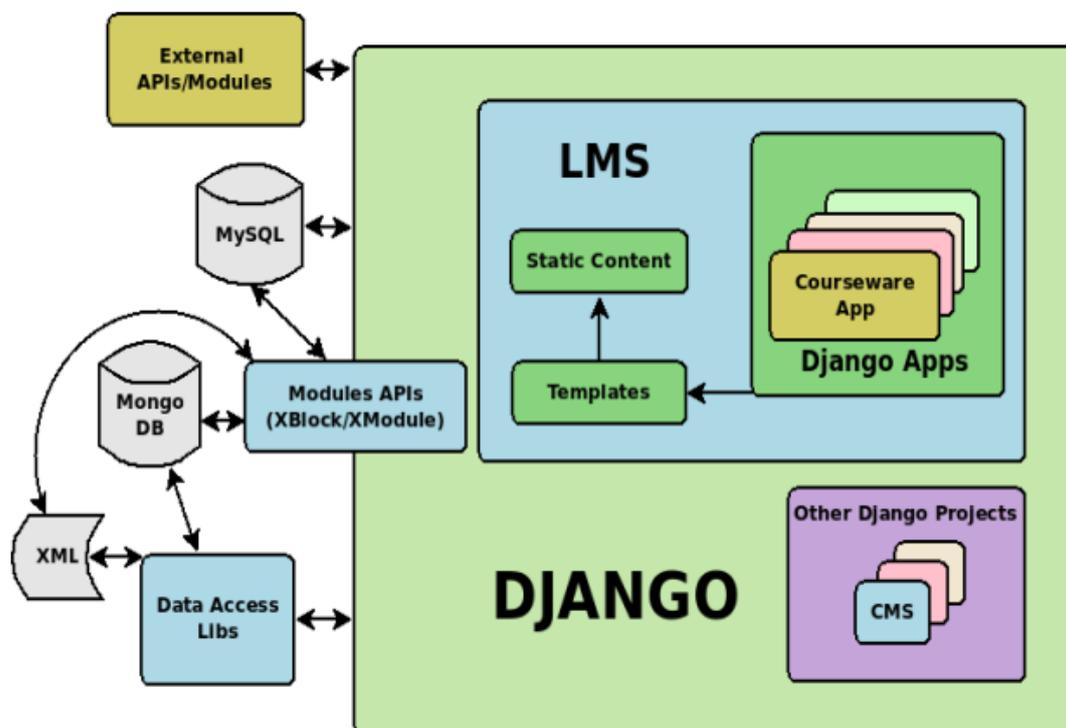


Figura 8. Base de datos en la arquitectura de Open edX.

Nota: Tomado de (Santofimia Ruiz, 2017)

3.4.2. Tracking Logs

Las acciones que realizan por los estudiantes y profesores dentro de la plataforma son recogidas por un log de eventos denominado tracking log. Los eventos recopilados por el archivo tracking log son emitidos por el servidor, el navegador o los dispositivos móviles; estos eventos recogen información relacionada a las interacciones con el material didáctico y los paneles del instructor dentro del LMS, los cuales son almacenados en documentos con formato JSON (edX, 2016).

Los eventos que son registrado en los trackings logs contienen campos comunes en su definición del esquema. Los campos que se encuentran en el nivel principal de los documentos JSON son: *accept_language*, *agent*, *context*, *event*, *event_source*, *event_type*, *host*, *ip*, *name*, *page*, *referer*, *sesión*, *time* y *username*, cada uno estos campos otorga información para conocer quien desencadeno el evento dentro de la plataforma (edX, 2016). En el **Anexo 1** se muestra un ejemplo de un evento registrado y anonimizado por la misma documentación oficial de edX y en la **Tabla 2** se da una breve descripción de cada campo guardado dentro de este archivo.

Tabla 2. Descripción de los campos que son registrados en el archivo tracking log.

Nota: Obtenido de (edX, 2016)

Campo	Tipo	Detalle
accept_language	string	Lenguaje en el que se dará la respuesta
agent	string	Agente del navegador de un usuario quien desencadenó el evento
context	Object	<p>Este campo contiene un conjunto de eventos que provén información de tipo textual y puede variar por el tipo de evento, pero sin embargo existen campos comunes en este campo como se muestra a continuación:</p> <p>course_id: Identifica el curso que generó el evento.</p> <p>org_id: la organización que proporcionó el curso</p> <p>path: la URL que generó el evento</p> <p>user_id: identifica al individuo que está realizando la acción, si un evento que se registra tiene un valor faltante o blanco, significa que un usuario cerró la sesión o la sesión se agotó mientras que la ventana del navegador permanecía todavía abierta.</p>
event	Object	Proporciona información específica de cada evento desencadenado, se proporcionan diferentes campos para diferentes eventos.
event_source	string	<p>Especifica el origen de la interacción que desencadenó el evento. Los posibles valores en este campo son:</p> <p>'browser'</p> <p>'mobile'</p>

UCUENCA

		'server' 'task'
event_type	string	El tipo de evento desencadenado. Los valores dependen de event_source
host	string	El sitio visitado por el usuario, por ejemplo, cursos.edx.org.
ip	string	Dirección IP del usuario que desencadenó el evento. Vacío para eventos que se originan en dispositivos móviles
name	string	Identifica el tipo de evento desencadenado
page	string	La '\$url' de la página que el usuario estaba visitando cuando se emitió el evento. Para los eventos de video que se originan en dispositivos móviles, identifica la URL del componente de video
referer	string	El URI del campo HTTP del encabezado de la solicitud
session	string	Este valor de 32 caracteres es una clave que identifica la sesión del usuario. Todos los eventos del navegador y los eventos de inscripción del servidor incluyen un valor para la sesión. Otros eventos de servidor y eventos móviles no incluyen un valor de sesión
time	string	Proporciona la hora UTC a la que se emitió el evento en formato 'AAAA-MM-DDThh: mm: ss.xxxxxx'.
username	string	El nombre de usuario del usuario que provocó la emisión del evento

3.5. Conclusiones

En este capítulo se presenta la arquitectura de la plataforma Open edX, sobre la cual se enfoca este trabajo de titulación. Los componentes que conforman la plataforma son diversos debido a que manejan una gran cantidad de información acerca de las interacciones de los estudiantes. La interfaz principal con la que interactúan los estudiantes se denomina LMS, la cual se encarga de publicar los MOOC y también brinda un sistema de administración a los profesores llamado Studio. Cada curso en Open edX está compuesto por módulos independientes llamados XBlocks, los cuales permiten proveer de funcionalidad a diferentes partes de un curso. Es así, que este trabajo de titulación se enfoca en el desarrollo de un componente XBlock que brinde soporte de LA a estudiantes y maestros dentro de la plataforma,

Open edX gestiona el registro de las interacciones de los estudiantes en archivos denominados tracking logs, los cuales son procesados posteriormente por nuestro componente de LA. Estos archivos contienen eventos relacionados al material didáctico de los cursos tales como el tiempo que dura la sesión de un estudiante, recursos específicos que son accedidos, fecha y otros metadatos que posteriormente ayudarán a entender las estrategias y patrones de aprendizaje de los mismos.

Capítulo 4: Análisis exploratorio del comportamiento de los estudiantes

4.1. Introducción

En este capítulo se ejecuta la primera y segunda etapa de la metodología de este trabajo de titulación, la cual comprende el análisis exploratorio de los datos y la identificación de variables. Primeramente, se realiza un análisis de los requerimientos del diseño de las visualizaciones. Se seleccionan aquellas tablas de la base de datos que aporten información relevante sobre el comportamiento de los estudiantes. Posterior a esto, se extraen y procesan los datos obtenidos de las interacciones de los estudiantes en el curso.

Una vez procesados los datos se identifican las variables relacionadas al éxito académico en entornos de aprendizaje virtual para luego analizar y seleccionar aquellas que mejor se adecuen a visualizaciones de analíticas de aprendizaje.

Finalmente, se analizan patrones y estrategias de aprendizaje con el fin de identificar los tipos de estudiantes que pueden existir dentro de un curso de aprendizaje virtual.

4.2. Análisis exploratorio de los datos

La exploración de los datos juega un papel importante al momento de analizar el comportamiento de los estudiantes (Purwoningsih et al., 2019b), ya que pueden llegar a convertirse en recomendaciones que mejoren las estrategias de aprendizaje de los mismos. Este paso también es conocido como EDA (Exploratory Data Analysis - por sus siglas en inglés). El uso del EDA en el área de LA permite alcanzar un mejor entendimiento acerca de los procesos de aprendizaje en los VLEs a través de factores como la cantidad de veces que un estudiante interactúa con los recursos de un curso, por poner un ejemplo.

En este trabajo de titulación se empleó el EDA y métodos de ML, específicamente clusterización, con el objetivo de analizar el comportamiento de los estudiantes, patrones y estrategias de aprendizaje que estos pudieran utilizar durante el transcurso de un curso dentro de la plataforma Open edX.

4.2.1. Análisis de requerimientos

El primer enfoque de este trabajo de titulación tiene como objetivo identificar variables e indicadores que permitan mejorar el éxito académico de los estudiantes en cursos MOOC. Por lo cual, como primer paso se debe realizar un análisis de requerimientos en el contexto de visualizaciones que den soporte a un dashboard de LA para la plataforma Open edX.

UCUENCA

El análisis de requerimientos permite identificar posibles características que pueden ser implementadas a través de las necesidades de los interesados del proyecto a desarrollar, en este caso los estudiantes de un curso. Martínez-Maldonado (2016), en su trabajo recomiendan que como primer paso se deberán abordar preguntas tales como:

- a) ¿Quiénes son los interesados?
- b) ¿Cuáles son las necesidades de los interesados?
- c) ¿Qué fuentes de datos se dispone?
- d) ¿Cuál es el contexto de aprendizaje?
- e) ¿Qué herramientas de LA se dispone?
- f) ¿Cómo están los datos siendo obtenidos, usados, compartidos y almacenados?

Tomando en cuenta lo anterior, para este trabajo de titulación se definen como interesados del proyecto aquellos estudiantes y maestros que deseen obtener información acerca del estado de su proceso de aprendizaje en un contexto de LA dentro de la plataforma Open edX. Los datos necesarios para cumplir los objetivos de este trabajo se encuentran almacenados en la base de datos de la misma plataforma, así como también en forma de logs que capturan las interacciones de los estudiantes al participar en curso. Estos datos serán posteriormente analizados, extraídos, procesados y almacenados por los autores de este trabajo de titulación haciendo uso del EDA con un script de Python, el mismo que estará disponible como anexo al final de este trabajo.

Con el fin de definir los requerimientos de estudiantes y maestros al momento de utilizar un dashboard de LA, se debe primero entender los objetivos de los mismos. Estos objetivos son un aspecto esencial del Aprendizaje Auto-Regulado (SRL - por sus siglas en inglés), en el cual se identifican las brechas entre el estado actual y el estado deseado de los estudiantes respecto a su proceso de aprendizaje en VLEs (Schunk & Zimmerman, 2012). Esto revela una necesidad del estudiante por (i) conocer cuál es situación actual dentro del curso, ya que esto le permite poder tomar decisiones para corregir o reforzar su proceso de aprendizaje. Otro aspecto importante que se revela es la necesidad de (ii) contrastar su rendimiento con otros compañeros dentro del mismo curso, esto como una motivación a mejorar o mantenerse dentro de aquel porcentaje de estudiantes que están teniendo éxito.

También es importante tomar en cuenta el nivel de compromiso que pueden adquirir los estudiantes al interactuar con los recursos dentro de un curso virtual. Este nivel de compromiso ha sido estudiado y organizado en tres dimensiones, tales como, medidas de comportamiento, medidas cognitiva y medidas

emocionales (Perez-Alvarez et al., 2020). Es así, que se evidencia el interés de estudiantes y maestros por (iii) conocer acerca de la frecuencia con la que son accedidas las lecturas, videos y problemas dentro de un curso. Además, de analizar la (iv) interacción de los estudiantes entre sí dentro de los foros del curso. Otro punto importante es el (v) tiempo invertido en las actividades, así como también el rendimiento, el cual es evidenciado a través de la (vi) cantidad de recursos que completan dentro de la plataforma.

4.2.2. Selección de tablas importantes de la base de datos

En esta subsección se introduce al lector de forma más detallada la estructura de la base de datos que utiliza Open edX para almacenar la información acerca de las interacciones de los estudiantes en un curso. Posteriormente, se seleccionan aquellas tablas que brinden información relevante acerca del estado de los estudiantes y su proceso de aprendizaje en la plataforma.

Tal como se explicó en la sección 3.3.1, la base de datos que maneja Open edX es MySQL, la cual se encarga de almacenar toda la información relacionada a las interacciones de los estudiantes en la plataforma. Los datos de los estudiantes se obtienen durante el registro en la plataforma y su posterior inscripción en algún curso, así como también las actividades que realicen con algún recurso en particular, tales como, un problema, lectura, video, etc.

Para este trabajo de titulación nos limitamos a analizar alrededor de las 305 tablas que dispone la base de datos denominada edxapp. De las cuales, se logró reducir a 22 tablas (ver **Anexo 2**) luego de un análisis detallado apoyándose en la documentación oficial de la plataforma (Open edX, n.d.).

Posterior al análisis inicial de la base de datos, se tomó como referencia trabajos anteriores (Javier et al., 2015; Khalil & Belokrys, 2020; Liang, Li, et al., 2016; Liang, Yang, et al., 2016; Pijeira Díaz et al., 2015; Ruipérez-Valiente, Muñoz-Merino, Pijeira Díaz, Santofimia Ruiz, et al., 2017; Ruiz et al., 2014; Volchek et al., 2017) relacionados a la plataforma Open edX, donde sus autores coincidían en utilizar las tablas que consideraban relevantes para obtener información acerca de las interacciones de los estudiantes. En la **Tabla 3** se detallan aquellas que fueron seleccionadas para este trabajo de titulación.

Tabla 3. Descripción de las tablas seleccionadas que contienen información relevante acerca de las interacciones de los estudiantes en la plataforma.

Tabla	Descripción
auth_user	Integración con el framework web de Open edX, la cual contiene información necesaria para el inicio de sesión de usuarios y permisos
student_courseenrollment	Cada fila en esta tabla representa la inscripción de un alumno en algún curso en particular. Se crea una fila para cada alumno que comienza el proceso de inscripción, incluso si nunca completa el registro en la plataforma activando la cuenta de usuario.
courseware_studentmodule	Contiene el estado más actual del curso, incluido el envío del problema más reciente y la unidad visitada en cada subsección.
xlea_indicators	Tabla elaborada por los autores. Esta tabla contiene los datos procesados de los trackings logs para ser visualizados en el dashboard de LA.

4.3. Identificación de variables para el mejoramiento del éxito estudiantil

En esta sección se presenta al lector las variables que servirán para identificar indicadores que mejoren el éxito académico de los estudiantes en los MOOC. Este análisis se basa en los estudios de trabajos relacionados a la mejora del rendimiento estudiantil en VLEs, de la mano de los requerimientos y datos relevantes obtenidos en la sección 4.1.

Varios autores definen al rendimiento estudiantil como la finalización de un examen final (Ramesh et al., 2014). También se define como alguna actividad realizada por los estudiantes en un punto de tiempo específico en el curso (Ramesh et al., 2014). Comúnmente, el análisis de las interacciones de los estudiantes es realizado por semanas. En algunos estudios, se cuestiona considerar como rendimiento estudiantil al simple de hecho de haber completado

un examen final (Clow, 2013; Koller et al., 2013), debido a que los estudiantes tienen diferentes motivaciones y objetivos de aprendizaje al momento de inscribirse en un MOOC.

Trabajos relacionados a la predicción del rendimiento estudiantil en MOOC se enfocan principalmente en la frecuencia con la que los estudiantes realizan actividades específicas (Conijn et al., 2018; P. G. de Barba et al., 2016; Ramesh et al., 2014). Los estudios revelan que los estudiantes más activos en un curso presentan una mayor probabilidad de mejorar su rendimiento académico en un MOOC. Además, se ha podido demostrar que las secuencias y patrones de aprendizaje de los estudiantes están relacionados a su éxito estudiantil (Brinton et al., 2015; Sinha et al., 2014); en estos trabajos se han utilizado técnicas de PM que evidencian que el orden de las actividades realizadas en un MOOC influyen en el rendimiento de los estudiantes.

Luego de analizar los resultados de los estudios mencionados anteriormente, en conjunto con los requerimientos y objetivos presentados en este trabajo de titulación, se abordarán como variables de éxito estudiantil a las siguientes:

- a) Frecuencia de actividades realizadas en el curso.
- b) Secuencias y patrones de aprendizaje de los estudiantes.
- c) Rendimiento del estudiante a través del tiempo.

4.3.1. Identificación de indicadores

Posterior al análisis realizado sección 4.2., en esta subsección se abordan indicadores ya explorados en estudios y trabajos previos relacionados al éxito académico de estudiantes en los MOOC, específicamente Open edX. Se presenta al lector una lista de posibles indicadores que den soporte a visualizaciones para un dashboard de LA.

A lo largo de varios estudios, se ha tratado como eje principal la necesidad de los estudiantes y sus expectativas al momento de desarrollar herramientas de LA en plataformas de aprendizaje virtual (West et al., 2020). Es a través de cuestionarios que investigadores del área han podido identificar indicadores que sirvan de soporte para dashboards de LA (Hilliger et al., 2020). Entre los hallazgos más relevantes se encuentra la necesidad de los estudiantes por conocer su situación en comparación de sus compañeros en el curso, con el objetivo de mejorar su proceso de aprendizaje y alcanzar el éxito académico (Bennett & Folley, 2020).

Tomando en cuenta lo anterior, se ha podido evidenciar que los estudiantes comúnmente eligen aquellos indicadores relacionados a la productividad y al

progreso del contenido en el curso (Conijn et al., 2018; Ioana, 2021). Estos indicadores están ligados a la idea de que completar actividades permite a los estudiantes motivarse a sí mismos a perseverar y finalizar el plan de estudio de un curso MOOC.

En otros trabajos relacionados directamente con la plataforma Open edX, los autores coinciden en que las frecuencias de actividades completadas, tales como números de lecturas, videos, cuestionarios y evaluaciones, son buenos indicadores debido a su fácil interpretación acerca del estado actual de un estudiante (Cobos et al., 2016; Javier et al., 2015; Khalil & Belokrys, 2020; Pijeira Díaz et al., 2015; Ruipérez-Valiente, Muñoz-Merino, Pijeira Díaz, Ruiz, et al., 2017; Ruiz et al., 2014). Otro aspecto importante a tomar en cuenta es el análisis de estos indicadores a través del tiempo (Breslow et al., 2013; Coffrin et al., 2014), comúnmente realizado en semanas, ya que permite a estudiantes y educadores entender el comportamiento progresivo dentro del curso y cómo afecta esto a su proceso de aprendizaje.

Posterior al análisis realizado en esta subsección, en la **Tabla 4** se presentan aquellos indicadores relacionados al éxito académico de estudiantes en plataformas de aprendizaje virtual.

Tabla 4. Indicadores relacionados al éxito académico en MOOC.

Indicador	Descripción
Revisión de contenido	Indica si los trabajos realizados por los estudiantes han sido revisados por un maestro o compañero
Participación en discusiones	La cantidad de veces que un estudiante interactúa en el foro.
Productividad	Cantidad de recursos completados en una sesión
Presencia en línea	El tiempo que un estudiante invierte dentro de la plataforma
Tiempo en el que comienza una actividad	La fecha en la que un estudiante inicia una actividad en el curso

UCUENCA

Tiempo para completar una actividad	La cantidad de tiempo que le toma a un estudiante completar una actividad
Actividades completadas	Cantidad de veces que un estudiante completa una actividad
Aportaciones en el foro	Retroalimentación positiva que reciben los estudiantes al participar en una discusión del foro
Videos completados	Cantidad de videos completados
Videos repetidos	Cantidad de veces que un estudiante ha repetido un video ya completado
Lecturas completadas	Cantidad de lecturas completadas
Lecturas repetidas	Cantidad de veces que un estudiante ha repetido una lectura ya completada
Exámenes completados	Cantidad de exámenes completados
Exámenes repetidos	Cantidad de veces que un estudiante ha repetido un examen ya completado
Cantidad de sesiones	Número de veces que un estudiante inicia sesión dentro del curso
Tiempo de sesiones	Duración de tiempo promedio que un estudiante invierte en una sesión
Nota actual del curso	Es el puntaje global que un estudiante ha obtenido durante todo el curso
Nota de examen	Es el puntaje obtenido por el estudiante luego de haber completado un examen
Adelantar video	Cantidad de veces que un estudiante a adelantado un video
Pausar video	Cantidad de veces que un estudiante a pausado un video

4.3.2. Análisis y selección de los indicadores

Finalmente, en esta subsección se identifican aquellas variables que mejoren el éxito académico de los estudiantes para luego ser implementadas en visualizaciones de LA. La selección de los indicadores se basa en los objetivos y requerimientos de este trabajo de titulación, además de los análisis realizados anteriormente.

Los indicadores mostrados en la **Tabla 5** se seleccionaron en base a las variables identificadas en la sección 4.2, las cuales se centran en la frecuencia de las actividades realizadas por los estudiantes, las secuencias y patrones de aprendizaje, y el rendimiento del estudiante a través del tiempo.

Tabla 5. Indicadores seleccionados para implementarse en el dashboard de LA.

Variable	Indicador	Descripción
Frecuencia de las actividades	interaction_foros	Cantidad de veces que un estudiante interactúa en los foros
	completed_lectures	Cantidad de lecturas completadas
	videos_complete	Cantidad de videos completados
	complete_problem	Cantidad de problemas completados
Secuencias y patrones de aprendizaje	repeated_lectures	Cantidad de veces que un estudiante repite una lectura
	repeat_videos	Cantidad de veces que un estudiante repite un video
	repeat_problem	Cantidad de veces que un estudiante repite un problema

UCUENCA

	num_play_videos	Cantidad de veces que se ha reproducido un video
	num_pause_videos	Cantidad de veces que se ha pausado un video
	num_stop_videos	Cantidad de veces que se ha detenido un video
	num_load_videos	Cantidad de veces que se ha accedido a un video
	num_fast_forward_videos	Cantidad de veces que se ha adelantado un video
	num_rewind_videos	Cantidad de veces que se ha retrocedido un video
	num_accessed_resources	Cantidad de recursos accedidos por un estudiante
Rendimiento a través del tiempo	week_id	Número de la semana en la que se desarrollan las actividades de un estudiante
	start_date_week	Fecha en la que comienza un periodo de actividades de 7 días
	end_date_week	Fecha en la que finaliza un periodo de actividades de 7 días
	num_sessions	Cantidad de veces que un estudiante se ha conectado a la plataforma.
	sum_time_sessions	Tiempo invertido del estudiante en la plataforma

4.4. Análisis del comportamiento de los estudiantes

En esta sección se procede a realizar un análisis relacionado al comportamiento y estrategias de aprendizaje que los estudiantes utilizan dentro de un curso de la plataforma Open edX. se informa al lector la estrategia utilizada en este trabajo de titulación para la extracción y procesamiento de los logs generados por los estudiantes en la plataforma. Para el análisis del comportamiento de los estudiantes se utilizaron técnicas de PM y ML, las cuales se detallan a lo largo de esta sección.

4.4.1. Extracción de los datos de la plataforma

Los datos que se usaron en este trabajo de titulación son datos de cursos impartidos por la universidad de Cuenca en el periodo 2020-2021 en su instancia de Open edX, en su versión Ironwood. En esta instancia de la plataforma se cuenta con cuatro cursos Fundamentos de Programación I, Fundamentos de Programación I - Parte II, Manejo inicial y prevención del trauma en niños y adolescentes y Ofimática; donde el número de inscritos varía entre 16 y 287 estudiantes. Entre estos cursos, el que se ha utilizado en este trabajo de titulación para el análisis es el curso de Ofimática que cuenta con 287 inscritos.

La instancia de la plataforma de Open edX usada en este trabajo de titulación es Koa por lo que se debe de realizar una migración de datos de Ironwood. La migración de los datos entre las dos versiones de la plataforma consta de nueve pasos que se mencionan a continuación: 1) tener instalado la última versión de Juniper en la plataforma; 2) Detener los servicios de la plataforma con el fin de que no genere conflictos en la importación de los datos; 3) realizar un backup de toda la información que se encuentra en la versión de Juniper, para ello es necesario emplear el script proporcionado por la propia documentación de la plataforma, que también se lo adjunta como **Anexo 3**; 4) copiar los archivos a la instancia de Koa; 5) detener todos los servicios de la instancia de Koa; 6) restaurar los datos de los datos de la instancia de Juniper a Koa, para ello es necesario usar los comandos de la propia documentación que también se lo adjunta en el mismo anexo; 7) ejecutar el proceso de la migración de los datos a Koa empleando el comando que se lo adjunta en el mismo anexo; 8) copiar los archivos de configuración de la maquina Juniper a Koa, dichos archivos se encuentra en el directorio de instalación y finalmente reiniciar los servicios para que se muestren los datos en la nueva plataforma. Adicionalmente es necesario copiar los archivos de los trackings logs de Ironwood a Juniper lo mismos que se encuentran en el directorio `'/edx/etc/log/tracking'`. Este procedimiento es crucial para tener todos los datos que se encuentran tanto en MySQL y MongoDB de

los cursos proporcionados por la Universidad de Cuenca en su plataforma de MOOC.

Los logs generados por la propia plataforma son de suma importancia para conocer el comportamiento de los estudiantes dentro de la plataforma y como se mencionó en el capítulo anterior estos son llamados tracking logs. En los trackings logs se registran una gran cantidad de eventos generados, pero no todos de ellos permiten conocer lo que está haciendo un estudiante en un momento específico, por lo que a lo largo de sección se muestran los eventos que aportan información sobre el comportamiento de un estudiante dentro de la plataforma.

Los videos en cualquier MOOC incluyen videos para explicar el contenido impartido, en esta plataforma no es la excepción y adicionalmente este cuenta con un su propio reproductor con controles familiares a muchos reproductores de video (edX, 2022). Este reproductor permite recolectar las acciones que los usuarios realizan al momento de observar un video y muchos de estos eventos son recolectados por el navegador web de las computadoras de escritorio o de los dispositivos móviles. En la **Tabla 6** se muestran los eventos recopilados que son almacenados en el archivo tracking logs referentes a la interacción de un usuario con un video de la plataforma (edX, 2016).

Tabla 6. Eventos registrados por los estudiantes cuando interactúan con la interacción con los videos.

Nota: Obtenido de (edX, 2016)

Evento	Breve Descripción
load_video/edx.video.loaded	Evento emitido cuando el video está completamente renderizado y listo para reproducirse, el navegador o la aplicación móvil.
pause_video/edx.video.paused	Cuando un usuario presiona el botón de pausa, el reproductor emite un evento pause_video. Para los videos que se transmiten en un navegador, cuando el reproductor llega al final del archivo de video y la reproducción se detiene automáticamente, emite tanto este evento como un evento stop_video

play_video/edx.video.played	Cuando un usuario selecciona el control de reproducción del reproductor de video
seek_video/edx.video.position.changed	Un navegador emite eventos seek_video cuando un usuario selecciona un control de interfaz de usuario para ir a un punto diferente en el archivo de video.
stop_video/edx.video.stopped	Cuando el reproductor de video llega al final del archivo de video y la reproducción se detiene automáticamente.

Además de los eventos generados por la interacción con los videos también existe otros eventos desencadenados por la interacción del usuario. Los otros eventos generados son relacionados con las evaluaciones realizadas tal como se muestra en la **Tabla 7**; interacción con los foros, mostrado en la **Tabla 8** y por último la interacción con los recursos de lectura donde el estudiante revisa instrucciones dejamos por el tutor o lecturas y estos están dentro de las unidades denominadas HTML (como se lo mencionó en el capítulo 3 de la estructura jerárquica de un curso) se muestran en la **Tabla 9**.

Tabla 7. Eventos registrados por los estudiantes con las evaluaciones.

Nota: Obtenido de (edX, 2016)

Evento	Breve Descripción
problem_check	Este evento se genera cuando el problema es revisado, este contiene información relacionada al id del problema, le número de intentos realizados, la calificación, la máxima calificación posible.
problem_reset	El navegador emite este evento de restablecimiento de problemas cuando un usuario hace clic en restablecer la respuesta.

Tabla 8. Eventos registrados por los estudiantes en los foros.

Nota: Obtenido de (edX, 2016)

Evento	Breve Descripción
edx.forum.comment.created	Evento emitido cuando los usuarios crean un comentario sobre una respuesta o dando sus contribuciones.
edx.forum.response.created	Los usuarios crean una respuesta a una publicación haciendo clic en Agregar una respuesta y luego enviando sus contribuciones.
edx.forum.response.voted	Los usuarios desencadenan este evento cuando dan clic en el icono de votar.
edx.forum.thread.created	Los usuarios crean un nuevo hilo también conocido como publicación y envían sus contribuciones.
edx.forum.thread.voted	Los usuarios indican un interés en un hilo creado anteriormente seleccionando el icono de votar.

Tabla 9. Eventos registrados con la interacción de recursos de lecturas.

Nota: Obtenido de (edX, 2016)

Evento	Breve Descripción
html_completed	Evento emitido cuando un usuario completa un html dentro de la plataforma.
html_repeated	Evento emitido por un usuario cuando después de completar un html lo vuelve a abrir.

Los eventos mencionados anteriormente permiten conocer lo que un estudiante realiza dentro de la plataforma, es por esto que para el desarrollo del script que recopila la información generada por los trackings log solo se ha considerado estos eventos para su análisis. El script esta realizado en el lenguaje de programación Python y el análisis de cada tipo de evento está dividido por funciones que analizan y extraen solo la información necesaria, entre esta

UCUENCA

información se encuentra el *user_id*, *course_id*, *block_id*, *event_type*, *time*, *category*. El campo *user_id* hace referencia al id del estudiante que está generando el evento; *course_id* id del curso en el que está el estudiante; *block_id* correspondiente al id del bloque que está analizando en ese momento; *event_type*, variable creada para poder identificar el tipo de evento que se está analizando y por último el campo *time* que permite conocer la fecha y hora que se desencadenó el evento. El código usado para analizar lo anteriormente mencionado se encuentra en el **Anexo 4**.

Una vez obtenido la información necesaria de los *trackings logs* se realiza un análisis a los datos generados por el estudiante en término de sesiones (registro de la actividad continua de un alumno dentro de la plataforma). En este trabajo se ha considerado el intervalo no mayor a 45 minutos pasado este tiempo se toma las actividades realizadas como otra sesión, esta definición de sesión es adoptada por los trabajos previos de (Kovanović et al., 2015) y (Boroujeni & Dillenbourg, 2018). El proceso de analizar las actividades continuas dentro de la plataforma y obtiene estadísticas en forma de sesiones se encuentran detallados en el **Anexo 5**.

Los MOOC estructuran sus contenidos en forma de una trayectoria lineal y las secuencias de interacción se puede estudiar siguiendo la trayectoria de aprendizaje que siguen los alumnos en función de la misma manera en que se estructura su contenido (Dan Davis et al., 2016). Con el objetivo de explorar la relación que existe entre el comportamiento de los estudiantes y su proceso de aprendizaje en los MOOC, es acertado realizar un análisis de sus diferentes actividades por semanas a lo largo del curso (van den Beemt et al., 2018a). De esta manera los datos obtenidos de las sesiones de los estudiantes son empleados para obtener estadísticas de las actividades realizadas a lo largo de la semana por los estudiantes. Además, se ha considerado que si un estudiante no termina la semana y se está ejecutando el script del análisis de información de los estudiantes se guarde la información no procesada en un archivo llamado '**data_daily_not_processed.csv**', esto con el fin de no perder información valiosa de un estudiante y que será usada en la próxima ejecución del script. Todo lo antes mencionado se puede evidenciar en código en el **Anexo 5**.

El procesamiento de los *trackings logs* hasta convertirlo en estadísticas que permitan mostrar en un dashboard como indicadores es almacenado en la base de datos propia de la plataforma, específicamente en MySQL. Las credenciales que permiten establecer una conexión con la base de datos se encuentran en el archivo *my-passwords.yml*, en el mismo directorio que se realizó la configuración de la instalación. Una vez establecido la conexión con la base de datos se almacena la información procesada de los *trackings logs* por medio de una tabla

UCUENCA

y un trigger para identificar el registro de la semana ingresada, la estructura con la que cuenta la tabla creada se muestra en la **Tabla 10**.

Tabla 10. Tabla creada para almacenar la información procesada de los trackings logs.

Campo	Tipo	Clave Primaria
user_id	int (11)	Si
course_id	varchar (50)	Si
week_id	int (11)	Si
start_date_week	timestamp	No
end_date_week	timestamp	No
num_sessions	int (11)	No
interaction_foros	int (11)	No
completed_lectures	int (11)	No
repeated_lectures	int (11)	No
videos_complete	int (11)	No
repeat_videos	int (11)	No
complete_problem	int (11)	No
repeat_problem	int (11)	No
sum_time_sessions	int (11)	No
num_play_videos	int (11)	No
num_pause_videos	int (11)	No
num_stop_videos	int (11)	No
num_load_videos	int (11)	No

num_fast_forward_videos	int (11)	No
num_rewind_videos	int (11)	No
num_accessed_resources	int (11)	No

4.4.2. Herramienta empleada para Minería de Procesos

El impacto que ha tenido la minería de procesos ha hecho posible que se desarrollen herramientas que permitan su uso automatizado. En el mercado existen una gran cantidad de herramientas que hacen posible la minería de procesos de una manera manejable, una de ellas es Disco (Dakic et al., 2019).

Disco es una herramienta construida por académicos con más de 8 años de experiencia en la minería de procesos, según los autores han usado esta experiencia para crear esta herramienta y que esta se adapte al flujo de trabajo de los profesionales, basándose en los tres pilares para su creación usabilidad, fidelidad y rendimiento (Fluxicon, 2022). A continuación, se a presentar las características más importante empleadas de este software:

- **Descubrimiento automatizado de procesos:** La tecnología que incorpora dentro de Disco permite crear mapas de procesos desde los datos sin ser procesados. Además, en los mapas generados se puede seleccionar el nivel de abstracción, visualizaciones de métricas directamente en el mapa y crear filtros a partir de las rutas o actividades (Fluxicon, 2022).

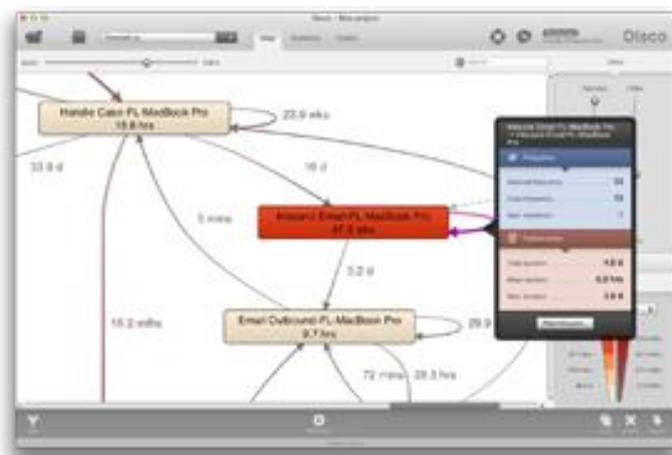


Figura 9. Descubrimiento automatizado de procesos.

Nota: Obtenido de (Fluxicon, 2022)

UCUENCA

- **Animación del mapa de procesos:** En el mapa de procesos se puede crear animaciones que sean visualmente atractivas, mostrando el proceso tal como sucedió y esto permite detectar los cuellos de botella de una manera instantánea observando donde se acumula el trabajo (Fluxicon, 2022).

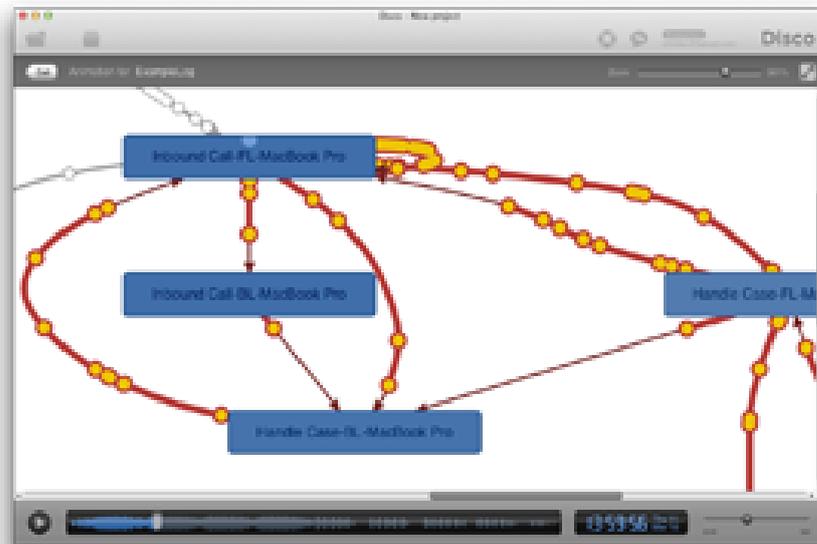


Figura 10. Animación del mapa de procesos.

Nota: Obtenido de (Fluxicon, 2022)

- **Estadísticas detalladas:** En el apartado de estadísticas disco se muestra información que ayuda a responder preguntas como: ¿Con qué frecuencia se ha realizado esta una actividad?, ¿Tiempo que se tarda en realizar las actividades?, entre otras. Con estas estadísticas se puede obtener una descripción general de los datos por medio de gráficos interactivos y tener información más detallada sobre una actividad, recurso y atributo (Fluxicon, 2022).

UCUENCA



Figura 11. Estadísticas detalladas de la herramienta.

Nota: Obtenido de (Fluxicon, 2022)

- **Casos:** En la opción de vista de casos, se puede inspeccionar el historial de los casos que son relevantes gracias a la función de búsqueda en vivo, además se puede apreciar los que siguen la ruta normal y cuales son excepciones (Fluxicon, 2022).



Figura 12. Vista de casos en Disco.

Nota: Obtenido de (Fluxicon, 2022)

4.4.3. Estrategias de aprendizaje e Identificación de patrones del comportamiento dentro del curso

Para responder a una pregunta de investigación, se emplea la adaptación de PM, PM2 de (Van Eck et al., 2015) realizada por Jorge Maldonado y sus colegas en (Maldonado et al., 2016). Las etapas que conforman esta adaptación son: Objetivos, Extracción de los datos, Generación del log de eventos y Descubrimiento y análisis del modelo. Estas etapas están representadas en forma gráfica en la **Figura 13**.

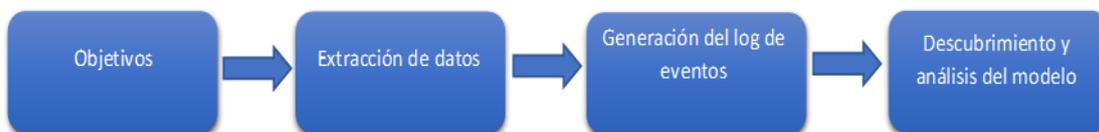


Figura 13. Etapas para generar el modelo de procesos.

Nota: Adaptado de (Van Eck et al., 2015)

4.4.3.1. Objetivos

- ¿Qué tipos de estudiantes podemos identificar al analizar su comportamiento dentro del curso?

4.4.3.2. Extracción de datos

En la plataforma de Open edX todas las interacciones de los estudiantes con los recursos del curso se almacenan en los archivos tracking logs que tienen una extensión json y son almacenados en la carpeta *'etc/var/log/tracking'* dentro de la plataforma. Una vez extraído los datos se selecciona solo los eventos que muestran el comportamiento de un estudiante dentro de la plataforma, los mismos que se muestran a continuación, *'load_video'*, *'play_video'*, *'stop_video'*, *'pause_video'*, *'seek_video'*, *'edx.video.position.changed'*, *'problem_check'*, *'edx.problem.hint.demandhint_displayed'*, *'problem_show'*; la información que contiene estos eventos y los campos de los eventos ya han sido detallados en este y el capítulo anterior.

La limpieza de los datos con los eventos seleccionados se realizó empleando el lenguaje de programación Python con la librería pandas. Se realizó de esta manera debido a que esta librería proporciona estructuras de datos rápidas, flexibles y expresivas que permite trabajar con conjuntos de datos estructurados que son comunes a las estadísticas, finanzas, ciencias sociales entre otros campos más (McKinney, 2018). Empleando esta librería se realizó la limpieza de los datos del id de los bloques para poder identificar qué actividad están realizando en ese momento, en el caso de los foros al no contar con un id del bloque se asignó el valor de vacío; los eventos de tipo *'seek_video'* y *'edx.video.position.changed'* no especifican si se avanzó o retrocedió en el video por lo que para este caso se analizó el tiempo de salto y se determinó si avanzo o retrocedió según sea el caso.

Después de realizar una limpieza de los datos se procedió a cambiar la etiqueta de la actividad realizada por una más descriptiva y fácil de analizar dentro de los datos obtenidos. En mucho de los casos estas etiquetas se han pasado al lenguaje y en el caso de los eventos que no especifican si se avanzó o no en el video se procedió a etiquetar con el valor de *'Adelanta video'*, o *'Retrocede video'* según sea el caso.

4.4.3.3. Generación del log de eventos

Después de la etapa de la extracción de los datos que se depuraron se muestra información detallada de como los estudiantes se fueron desarrollando en el curso a lo largo del tiempo dentro de la plataforma de Open edX. La información anteriormente recopilada servirá para responder la pregunta de investigación planteada.

UCUENCA

Para responder a la pregunta de investigación se generó un archivo con formato csv con la información necesaria para poder develar que tipo de comportamiento y estrategias tienen los estudiantes dentro del curso, las columnas basadas en la variable 'secuencia y patrones de aprendizaje' que muestren el comportamiento de los estudiantes en el curso son las siguientes:

- **num_completed_lectures:** Contiene el número de lecturas que ha completado a lo largo del curso un estudiante.
- **num_repeated_lectures:** Número de lecturas que el estudiante ha revisado nuevamente después de completarlas.
- **num_videos_complete:** Contiene el número de video que ha completado un estudiante dentro del curso.
- **num_repeat_videos:** Número de lecturas que el estudiante ha vuelto a revisar después de haberlas finalizado.
- **num_play_videos:** Las veces que un estudiante ha dado click en el botón de reproducir un video.
- **num_pause_videos:** Las veces que un estudiante ha dado click en el botón de pausar un video.
- **num_stop_videos:** Las veces que un estudiante ha dado click en el botón de detener un video.
- **num_complete_problem:** Número de problemas completados por el estudiante, se debe de mencionar que una evaluación está construida por numerosos problemas.
- **num_repeat_problem:** El número de problemas que ha repetido un estudiante, se debe de mencionar el que número de intentos que tiene un estudiante lo determina el tutor que está impartiendo el curso.
- **num_load_videos:** El número de veces que un estudiante carga un video, pero esto no significa que vaya a reproducir dicho video.
- **sum_time_sessions:** El tiempo sumado de las sesiones de un estudiante a lo largo del curso.

Finalmente, se procedió a utilizar la técnica de ML para agrupar a los estudiantes según sus características. Para este proceso se empleó el algoritmo k-means, uno de los algoritmos más relevantes que se tiene en el clustering, el cual tiene como objetivo encontrar y agrupar los datos que tienen similitud entre ellos, cuando más cerca estén los puntos más probabilidad de pertenecer al mismo cluster. Este algoritmo se caracteriza por usarse en un conjunto de datos que no se encuentra etiquetados es decir no tienen grupos previamente definidos (Zhang, 2020). En el clustering los parámetros de entrada son importantes es por ellos que existen técnicas e índices para su validación (Congming Shi et al., 2021). Más adelante se muestra el proceso empleado para la validación en el clustering.

4.4.3.4. Descubrimiento y análisis de los modelos

Después de obtener el log de eventos se realizó un análisis para el descubrimiento de los modelos con el propósito de emplearlos y responder a la pregunta de investigación “¿Qué tipos de estudiantes podemos identificar al analizar su comportamiento y estrategias de aprendizaje dentro del curso?”. En este sentido se realizó un enfoque en base a las actividades que realizó un estudiante, las cuales se armaron con variables que no se encuentran dentro de los logs de eventos de Open edX de manera explícita, se tuvieron que calcular en base a la interacción de los estudiantes con la plataforma. Para todos los estudiantes del curso se calcularon las variables: a) número de lecturas completadas, b) número de lecturas repetidas, c) número de videos completados, d) número de vides repetidos, e) número de veces que ha dado clic en el botón de reproducir, f) número de veces que ha dado clic en el botón de pausar, g) número de veces que ha dado clic en el botón de detener, h) número de veces que ha completado un problema, i) número de problemas que ha repetido el estudiante, j) número de veces que ha cargado el video (esto no significa que se reproduzca automáticamente el video), k) el tiempo que ha empleado el estudiante en todas las sesiones en la plataforma. Esto se realizó de esta manera para conocer como es el comportamiento de los estudiantes en la plataforma y las estrategias que estos emplean, esto se lo muestra en la **Figura 14**.

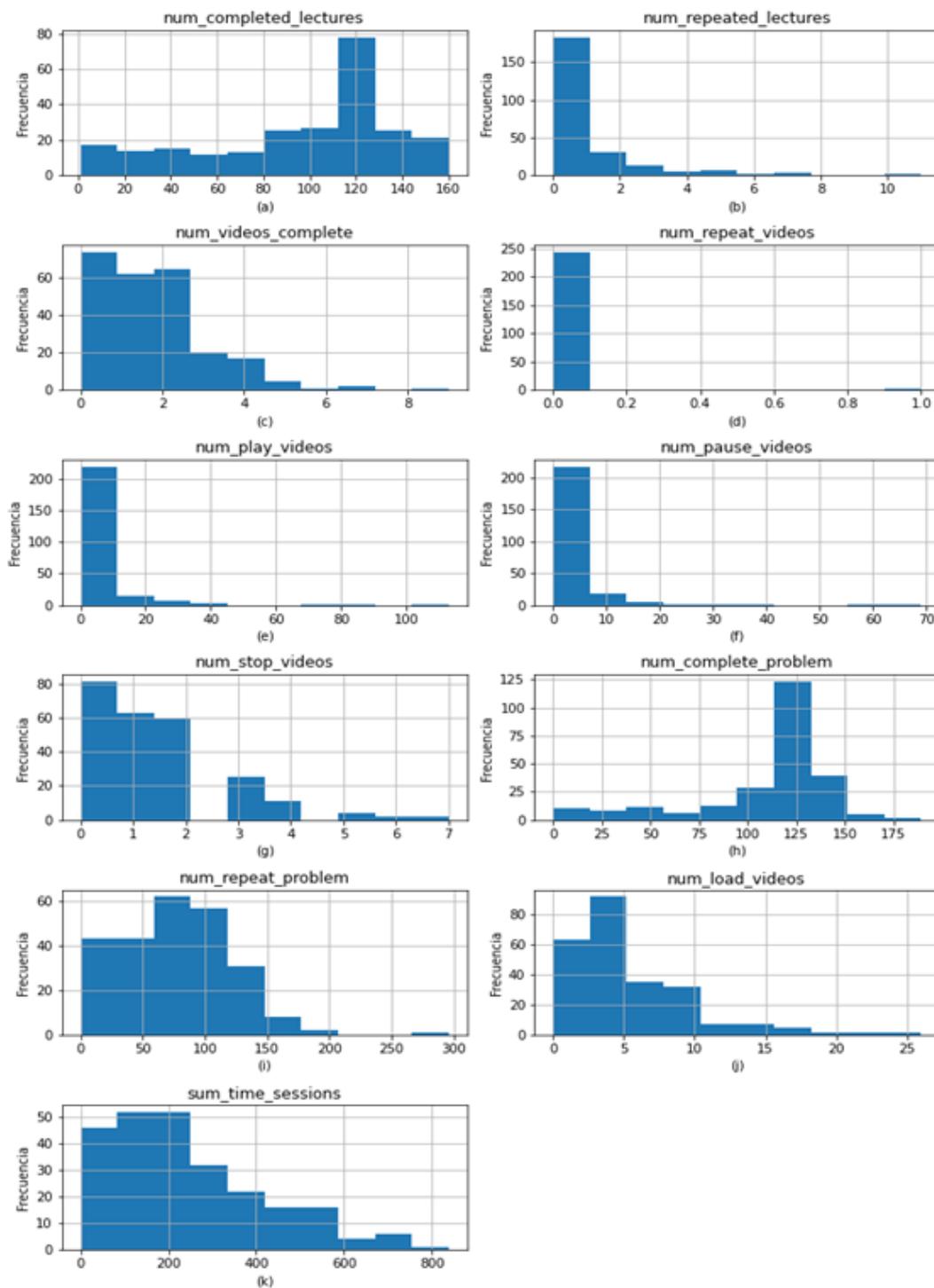


Figura 14. Distribución de las variables para caracterizar los grupos del cluster.

Después de realizar la identificación de las variables para el agrupamiento de los estudiantes, se empleó la técnica de ML para agrupar a los estudiantes según sus características. El algoritmo empleado es para realizar el agrupamiento es k-means; para obtener el número de clusters resultantes, se realizó una validación previa de dos maneras, la primera es de manera cualitativa empleando el método del codo (Elbow Method en inglés) se caracteriza por graficar la suma del cuadrado de las distancias de cada objeto del cluster a su centroide, con al grafico generado se busca encontrar un cambio de pendiente (un codo) para obtener el número de más idóneo de clusters (Et-Taleby et al., 2020); La segunda es de manera cuantitativa empleando el método de la silueta (Silhouette Method en inglés) el cual se basa en mostrar la separación de los grupos resultantes, este grafico muestra que tan cerca este cada punto de un cluster de los clusters vecinos, un valor de un coeficiente de silueta cercano a 1 significa que el grupo este lejos de los grupos vecinos, un valor de 0 esta entre dos grupos vecinos y finalmente un valor de -1 puede interpretarse como que se ha asignado a un grupo equivocado; por lo que un valor cercano a 1 nos da un mejor valor de los clusters (Jan Štrobl et al., 2017). Empleando ambos métodos de validación se obtuvo que tener un valor de 3 clusters es el idóneo para nuestro análisis, tal como se lo muestra en la **Figura 15 y 16**.



Figura 15. Método del codo para determinar el número ideal de grupos.



Figura 16. Método de la silueta para determinar el número ideal de grupos.

Visualizar los clusters de manera gráfica permiten observar cómo están agrupándose los datos, por lo que se requiere llevar de a un espacio de dos o tres dimensiones. La reducción de dimensionalidad permite mantener la relación original de los datos de alta dimensión a espacios de dimensionalidad reducida sin perder su relación, lo cual es útil para el preprocesamiento o el análisis visual de datos de alta dimensionalidad (Gong et al., 2020). Las técnicas de reducción de dimensionalidad comúnmente usadas son el análisis de componentes principales (PCA por sus siglas en inglés) y t-SNE (t-stochastic neighbor embedding) (Gong et al., 2020). Empleando ambas técnicas de reducción de dimensionalidad con el propósito de reducir a solo dos dimensiones para poder graficar se obtuvo los resultados que se muestran en las **Figuras 17 - 18**, donde se observa los tres clusters cada uno con un diferente color.

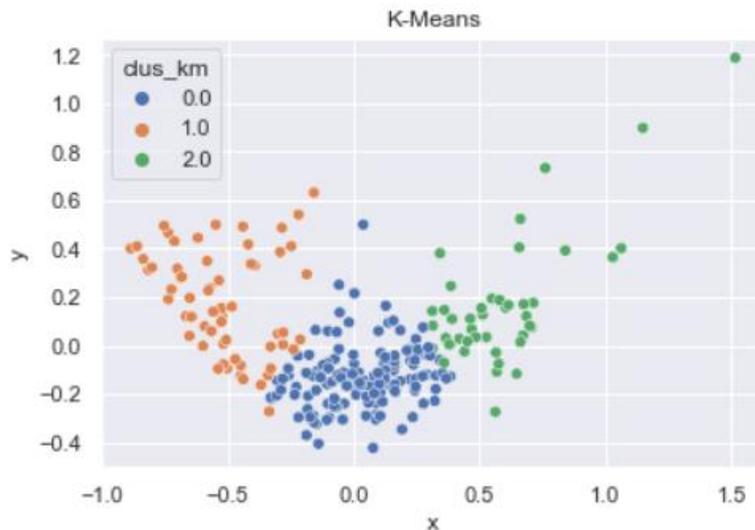


Figura 17. Visualización obtenida empleando el método de reducción de dimensionalidad PCA.

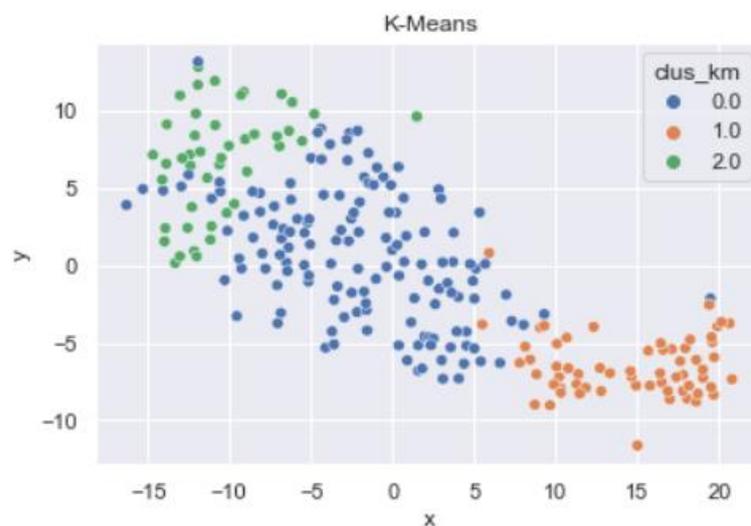


Figura 18. Visualización obtenida empleando el método de reducción de dimensionalidad t-SNE.

Con el propósito de responder a la primera parte de la pregunta de investigación referente a los tipos de estudiantes que podemos identificar, es necesario realizar un análisis a los clusters obtenidos. Una medida que nos permite conocer como es el comportamiento de todos los estudiantes es la media de cada variable analizada, tal como se lo muestra en la **Tabla 11**. En esta tabla

se muestra una variable denominada interacción con los videos la misma que fue el resultado de todos los eventos con los que el estudiante puede interactuar en un video, de igual manera las lecturas repetidas y completadas se encuentran como lecturas completadas y de esta misma manera se realizó con los videos para tener un mejor entendimiento al análisis que se va a realizar.

Tabla 11. Analisis usando la media de los clusters obtenidos.

Descripción	Cluster 0	Cluster 1	Cluster 2
Número de estudiantes	141	62	44
Lecturas completadas	115.48	34.01	134.93
Interacción con los videos	8.65	5.12	41.29
Videos completados	1.24	0.53	3.84
Problemas completados	200.14	116.67	216.32
Videos cargados	4.69	2.93	11.27
Tiempo de las sesiones	261.74	83.77	427.57

Empleando un análisis a los valores obtenidos en la Tabla 11, se puede identificar a los siguientes tipos de estudiantes dentro de la plataforma. En el cluster 0 se tiene a los estudiantes que evidencian una cantidad de problemas completados inferiores al cluster 2. Sin embargo en este grupo también se encuentran aquellos estudiantes que en promedio no han completado videos ni tantas lecturas como el cluster 2, por consiguiente este grupo cae en la definición de estudiantes estratégicos u objetivos (Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, et al., 2018). Este tipo de estudiantes tienden a centrar sus esfuerzos en pasar las evaluaciones a fin de certificar o probar sus conocimientos.

En el cluster 1, se observan características que definen a este grupo como estudiantes de tipo muestreo (Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, et al., 2018). Este grupo está compuesto por un promedio de estudiantes que tiene un tiempo de sesiones menor a los otros grupos, además el número de lecturas completadas, videos completados y problemas completados es menor a los otros clusters; estos alumnos son aquellos que no completan las evaluaciones ni problemas propuestos, y las interacciones con los recursos del curso son erráticas.

UCUENCA

Finalmente en el cluster 2, se identifica a los estudiantes comprensivos (Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, et al., 2018). Este grupo se encuentra compuesto por aquellos estudiantes que más videos han completado a comparación de los otros clusters. Además, estos estudiantes son los que más veces han interactúa con el mismo video y por lo tanto exhiben un mayor nivel de compromiso para terminar con éxito el curso.

Después de revisar las diferencias entre la media de las estadísticas se emplea el algoritmo de la minería de procesos con la ayuda de la herramienta Disco, con el propósito de responder la segunda parte de la pregunta de investigación, la misma que es conocer las estrategias de aprendizaje dentro del curso. Disco permite visualizar como es el comportamiento de los estudiantes empleando un algoritmo de PM y poder identificar las secuencias de aprendizaje que emplean los estudiantes dentro de cada cluster. En la **Figura 19** se muestra el mapa de las actividades que han tenido los estudiantes dentro del cluster 0, y de igual manera las actividades de los cluster 1 y 2 se muestran en las **Figuras 20-21** respectivamente.

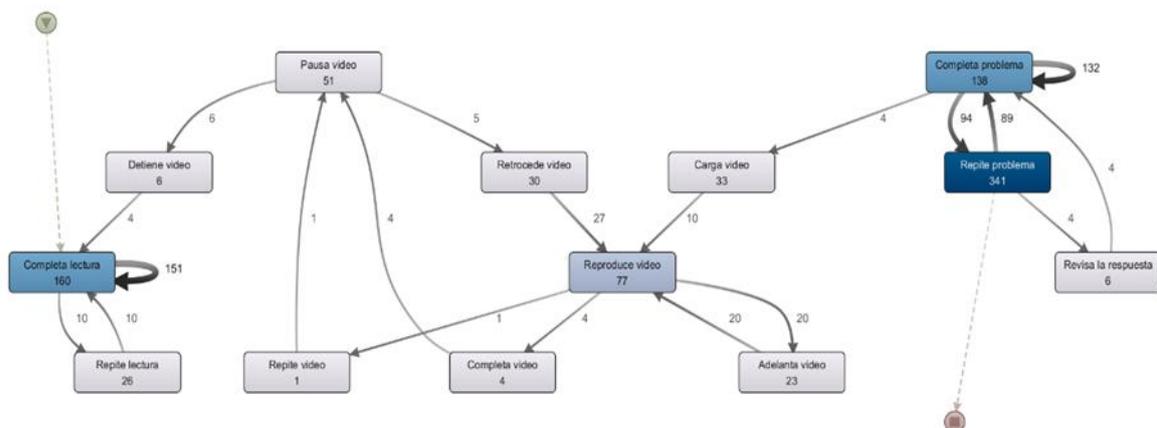


Figura 19. Modelo de procesos para el Clúster 0.

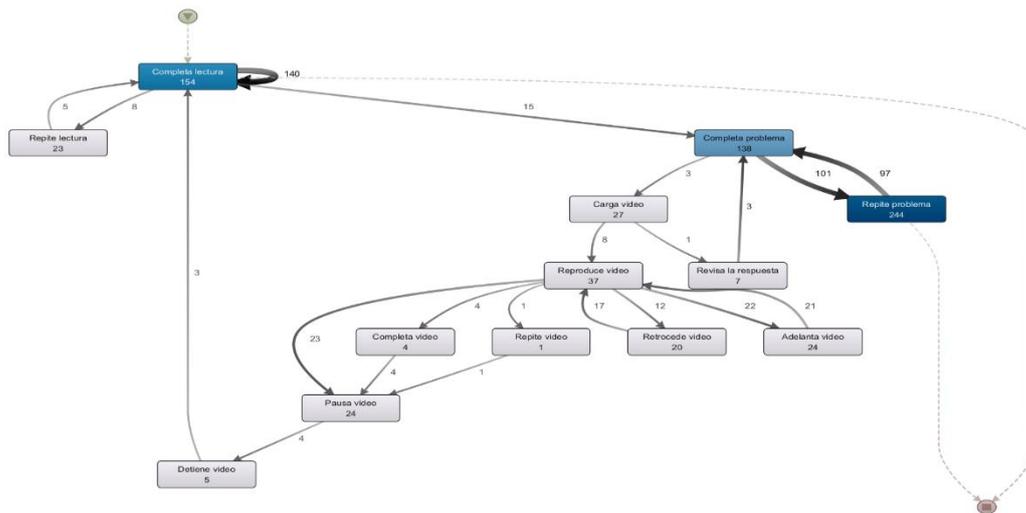


Figura 20. Modelo de procesos para el Clúster 1

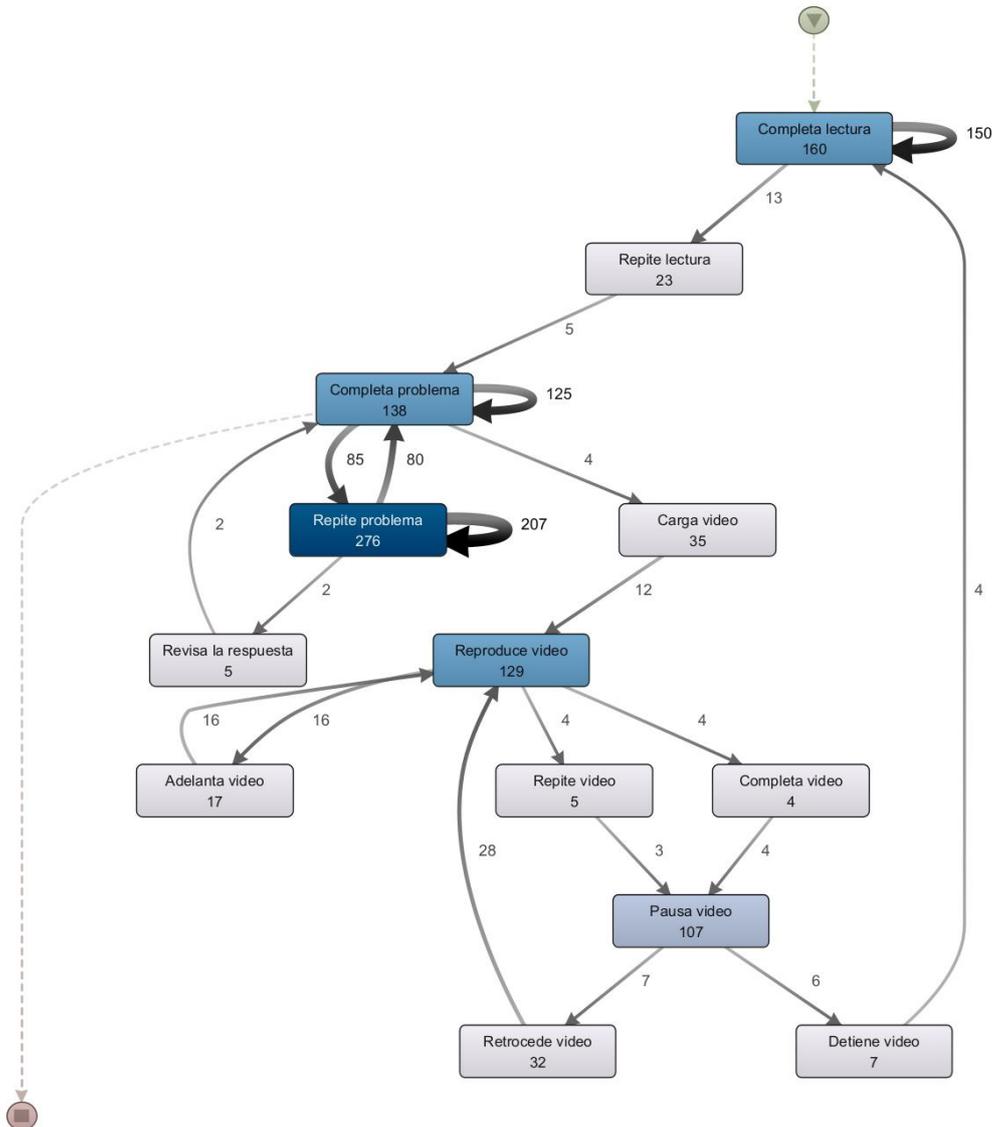


Figura 21. Modelo de procesos para el Clúster 2.

Después de obtener las secuencias de aprendizaje de los estudiantes en cada uno de los clusters, se analizó cada una de las secuencias de aprendizaje a un nivel macro, en este análisis se pudo identificar las siguientes secuencias:

Cluster 0:

- a) Completa lectura → Completa lectura

UCUENCA

- b) Completa lectura → Repite lectura → Completa lectura
- c) Completa problema → Completa problema
- d) Completa problema → Repite problema → Completa problema
- e) Completa problema → Revisa la respuesta → Repite problema → Completa problema
- f) Completa problema → Carga video → Reproduce video → Adelanta video → Reproduce video → Completa video → Pausa video → Detiene video → Completa lectura

De las secuencias de aprendizaje que se encuentran el cluster 0 se puede decir que estos estudiantes tienen un comportamiento de ingresar al curso completar los problemas y cuando terminan la evaluación van a revisar los videos y las lecturas como una manera de poder revisar el contenido y encontrar la respuesta a los problemas que no pudieron responder.

Cluster 1:

- a) Completa lectura → Completa lectura
- b) Completa lectura → Repite lectura → Completa lectura
- c) Completa lectura → Completa problema → Repite Problema
- d) Completa lectura → Completa problema → Repite Problema → Carga video → Reproduce video → Pausa video → Detiene video → Completa lectura
- e) Completa lectura → Completa problema → Repite Problema → Carga video → Revisa la respuesta → Completa problema → Carga video → Reproduce video → Retrocede video → Reproduce video → Completa video → Pausa video → Detiene video → Completa lectura

De las secuencias de aprendizaje que se encuentran el cluster 1 y de su análisis se puede decir que estos estudiantes tienen un comportamiento de ingresar a revisar las lecturas del curso, intentan completar algunos problemas y finalizan la sesión, muchos de ellos solo revisan el contenido del curso y finalizan la sesión.

Cluster 2:

- a) Completa lectura → Repite lectura → Completa lectura
- b) Completa lectura → Completa lectura
- c) Completa problema → Completa problema
- d) Completo problema → Repite problema
- e) Completa lectura → Repite lectura → Completa problema
- f) Completa lectura → Repite lectura → Repite Problema → Carga video → Adelanta video → Reproduce video → Completa video → Pausa

video → Reproduce video → Repite video → Pausa video → Detiene video → Completa lectura → Repite lectura → Completa problema

De las secuencias de aprendizaje que se encuentran el cluster 2 se puede decir que una vez que completan las lecturas vuelven a revisarlas para dar paso a responder las preguntas de las evaluaciones y en caso de que no hayan pasado las evaluaciones revisan de nuevo el material del curso tanto las lecturas como los videos para volver a dar las evaluaciones.

De los tres clusters, las secuencias que más se repite en el comportamiento de los estudiantes es la revisión de lecturas e inclusive repetir las mismas. De igual manera este comportamiento también se visualiza en los problemas, sin embargo, este no es el caso de los videos ya que existe una baja cantidad de estudiantes que los completan e inclusive muchos ni siquiera los vuelven a revisar.

4.5. Conclusiones

Este capítulo aborda la primera etapa de la metodología de este trabajo de titulación. La exploración de los datos juega un papel importante al momento de analizar el comportamiento de los estudiantes, es debido a esto que para el desarrollo de nuestro componente de LA se empleó técnicas como EDA y métodos de ML con el objetivo de obtener información acerca del comportamiento de los estudiantes, patrones y estrategias de aprendizaje. Para la selección de las tablas importantes de la base de datos, previamente se realizó un levantamiento de requerimientos. Se analizaron alrededor de 305 de las que almacena Open edX en su sistema, de las cuales se consideraron aquellas que brinden información acerca de los cursos y las interacciones de las estudiantes, tales como *auth_user*, *student_courseenrollment*, *courseware_studentmodule*, y la finalmente la tabla *xlea_indicators* generada por nuestro componente de LA.

Posteriormente se analizaron e identificaron indicadores relacionados al éxito académico de estudiantes en MOOC, con el objetivo de incorporarlos en las visualizaciones del componente desarrollado en este trabajo de titulación. Los indicadores seleccionados se centran en la frecuencia de las actividades realizadas por los estudiantes, las secuencias y patrones de aprendizaje, y el rendimiento del estudiante a través del tiempo.

Finalmente, se realizó Minería de Procesos con la herramienta Disco sobre los datos procesados de los estudiantes con el fin de descubrir patrones y secuencias de aprendizaje que ayuden a entender el comportamiento de los estudiantes dentro de la plataforma Open edX. De este análisis se obtuvieron tres grupos de estudiantes, los cuales se identificaron como estudiantes

UCUENCA

estratégicos, estudiantes de nuestro y estudiantes comprensivos. Las secuencias que más se repiten en el comportamiento de los estudiantes es la revisión de lecturas, así como su repetición luego de completarlas. De igual manera se evidenció comportamiento similar relacionado a cuestionarios y exámenes, sin embargo, este no es el caso de los recursos audio visuales, donde existe una baja cantidad de estudiantes que los completan e incluso ni los revisan.

Capítulo 5: Diseño de las visualizaciones

5.1. Introducción

Posterior al análisis realizado y los indicadores obtenidos previamente, en este capítulo se ejecuta el tercer paso de la metodología la cual comprende la planificación y diseño empleada en este trabajo de titulación. Se realiza una revisión exhaustiva sobre visualizaciones ya existentes que aporten LA a los MOOC. El propósito de este capítulo es seleccionar aquellas visualizaciones que cumplan con los objetivos y requerimientos expuestos anteriormente, para luego adaptarlas a un componente de LA para la plataforma Open edX.

5.2. Revisión y selección de visualizaciones existentes

Open edX en su objetivo de brindar un dashboard de analíticas acerca del estado de la plataforma y los estudiantes que se registran a sus cursos, desarrolló y liberó su propio dashboard llamado Insights (*EdX Insights*, n.d.). EdX Insights permite que la información de cada curso esté disponible para los administradores de la plataforma, proporcionando datos acerca de la actividad, los antecedentes y rendimiento de los alumnos a lo largo de un curso (edX, 2018). A pesar de esto, las visualizaciones que brinda Insights carecen de información específica acerca de cada estudiante (ver **Figura 22**), además de solo estar disponible para los administradores de la plataforma.

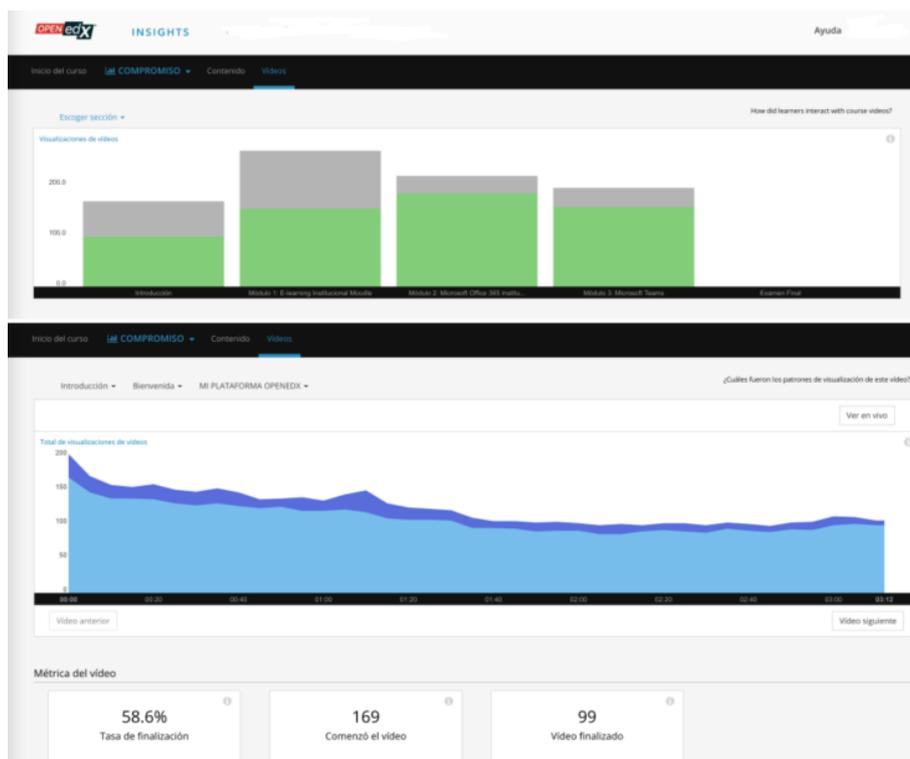


Figura 22. Dashboard de edX Insights para medir el nivel de compromiso de los estudiantes dentro de un curso.

Nota: Obtenido de (EdX Insights, n.d.)

Debido a esto, proyectos como ANALIZE y OXALIC (Khalil & Belokry, 2020; Ruipérez-Valiente, Muñoz-Merino, Pijera Díaz, Ruiz, et al., 2017), han desarrollado sus propios dashboards de LA para ser utilizados en cursos de Open edX. Estas herramientas se enfocan en el proceso de aprendizaje de los estudiantes dentro de un curso, para luego mostrar esta información a los profesores mediante un pequeño dashboard.

ANALYZE incluye en su módulo de LA, visualizaciones acerca del progreso de las actividades dentro del curso. En la **Figura 23** se muestra un diagrama de barras aplicado a recursos evaluables dentro del curso, es decir, tareas y exámenes. Esta visualización muestra el porcentaje de estudiantes que han completado exitosamente, aprobado y fallado algún recurso evaluable.

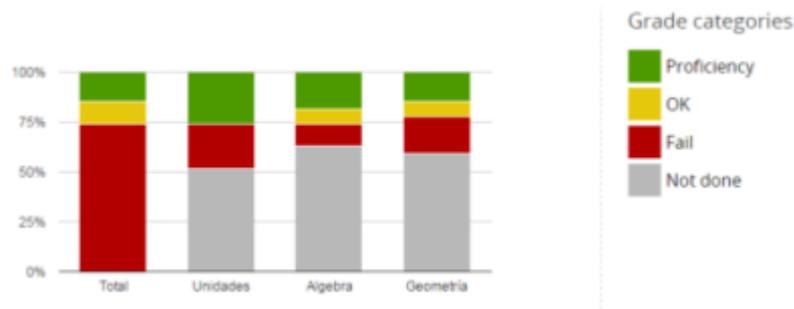


Figura 23. Visualización del resumen de un curso.

Nota: Tomado de (Santofimia Ruiz, 2017).

En la **Figura 24** se muestra una visualización relacionada a la actividad de los estudiantes en el curso. Se visualiza un diagrama de pie en el que se muestra el tiempo invertido por cada estudiante con los capítulos de un curso.

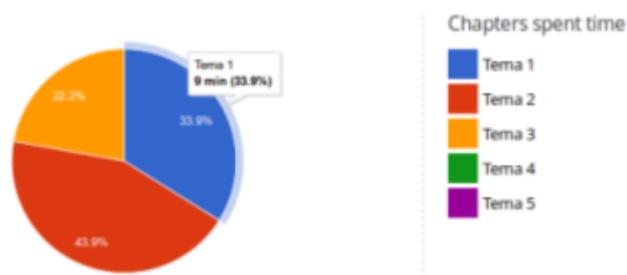


Figura 24. Visualización del tiempo invertido en cada capítulo.

Nota: Tomado de (Santofimia Ruiz, 2017).

ANALYZE también cuenta con visualizaciones que permiten observar la cantidad de veces que un estudiante ha accedido algún capítulo dentro del curso (ver **Figura 25**). Se visualiza un gráfico de barras con los capítulos de un solo curso, en el cual se indica el promedio de accesos a cada uno.

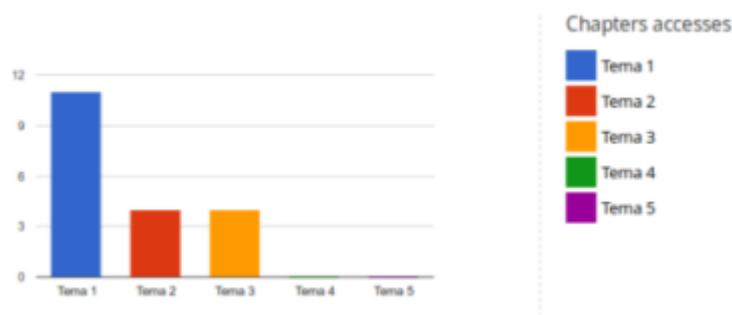


Figura 25. Visualización de la cantidad de accesos de cada capítulo.

Nota: Tomado de (Santofimia Ruiz, 2017).

Finalmente, ANALYZE muestra una visualización acerca del progreso del estudiante en los videos y las notas obtenidas en los problemas (ver **Figura 26**). Se visualiza una gráfica de área a través del tiempo en el curso.

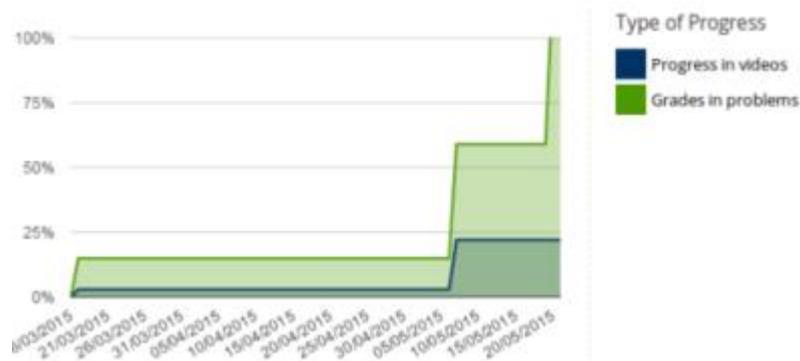


Figura 26. Visualización del progreso de videos y problemas.

Nota: Tomado de (Santofimia Ruiz, 2017).

Por su parte, OXALIC presenta sus visualizaciones en un dashboard compacto que permite a los usuarios interactuar de manera más sencilla con la información presentada. En la **Figura 27** se muestra la organización de cada gráfica dentro del componente de LA de OXALIC.



Figura 27. (a) Menu principal de OXALIC. (b) Resumen del número total de interacciones con los vídeos. (c) Número de interacciones para cada vídeo. (d) Distribución de eventos relacionados a la reproducción de los vídeos. (e) Tiempo en la plataforma, inscripciones, sesiones y duración.

Nota: Tomado de (Khalil & Belokry, 2020).

5.3. Criterios para la selección de las visualizaciones

Open edX y las herramientas mencionadas anteriormente siguen pautas establecidas para introducir a los usuarios de un dashboard de LA (edX, 2018). Estas pautas se basan en:

UCUENCA

- Visualizar métricas acerca la inscripción de los estudiantes.
- Seguimiento de actividades
- Compromiso con el contenido del curso
- Compromiso con videos del curso

Basándose en estas pautas, y acorde a los objetivos, requerimientos e indicadores expuestos en el capítulo anterior, se hizo uso de la guía para la elección de visualizaciones para dashboards de aprendizaje “Linking dashboard design and data visualization concepts” (Sedrakyan et al., 2019). Esta guía y la metodología utilizada en este trabajo de titulación (Martinez-Maldonado et al., 2016) se enfocan en los siguientes aspectos al momento de elegir dashboards de visualizaciones para LA:

- a) Objetivos y necesidades de los interesados
- b) Fuente de los datos
- c) Retroalimentación y autorregulación del aprendizaje
- d) Facilidad para medir el proceso de aprendizaje

Estudiantes y maestros tienen la necesidad de a) conocer acerca de la frecuencia con la que se interactúa con los recursos de la plataforma. Además, requieren analizar su comportamiento y tiempo invertido en semanas. Los b) datos obtenidos para alimentar las visualizaciones de LA son procesados en base a los indicadores obtenidos en el capítulo 4.

Una de las estrategias de c) SRL comúnmente utilizadas por este tipo de herramientas de LA es identificar las brechas entre el estado actual y el estado deseado de los estudiantes con respecto al promedio de sus compañeros dentro de un curso (Schunk & Zimmerman, 2012). Con el objetivo de d) facilitar la interpretación del estado pasado y actual del proceso de aprendizaje las visualizaciones deben estar desglosadas en periodos de tiempo uniforme, comúnmente dividido en semanas (Ramesh et al., 2014).

5.4. Diseño de las visualizaciones

Finalmente, en esta sección se presenta al lector un prototipo de las visualizaciones a ser implementadas en el componente de LA para Open edX. Este prototipo se diseñó en base a los aspectos obtenidos en la sección 5.2., los cuales se enfocan en brindar a estudiantes y docentes gráficas que detallen el estado del proceso de aprendizaje a través del tiempo dentro de un curso.

En la **Figura 28**, se muestra el prototipo que será presentado a los estudiantes dentro de un curso. Este prototipo contiene visualizaciones acerca de i) la cantidad de recursos completados por semanas, ii) interacciones con los videos

del curso, iii) total de lecturas, problemas y videos completados a lo largo de todo el curso y el iv) tiempo invertido. La mayoría de estas visualizaciones se comparan con el promedio del curso, con el objetivo de dar una retroalimentación al estudiante para que pueda autorregular su aprendizaje a través de los datos.

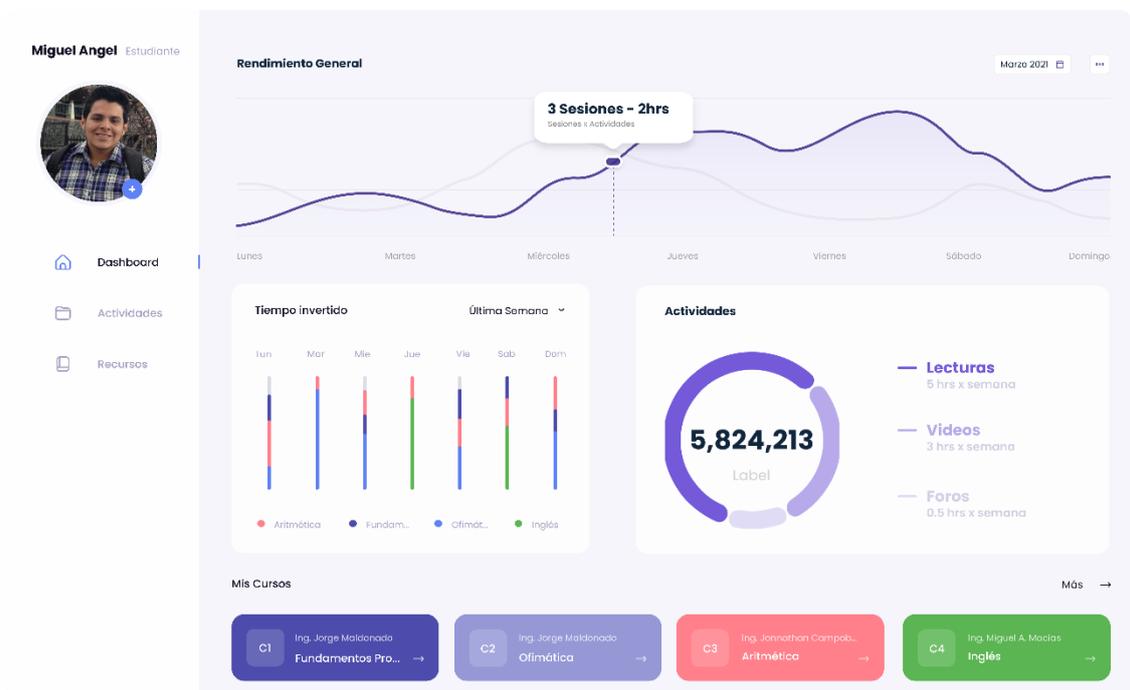


Figura 28. Prototipo del dashboard de la vista del estudiante.

Para el prototipo de la vista del docente (ver **Figura 29**) se presentan visualizaciones acerca de los i) recursos completados por semanas, ii) interacciones con videos, iii) sesiones en promedio por semana, iv) recursos accedidos en promedio por semana y v) una tabla con estadísticas acerca de cada estudiante por individual. Estas visualizaciones se obtienen a través del promedio de todos los estudiantes semana a semana.

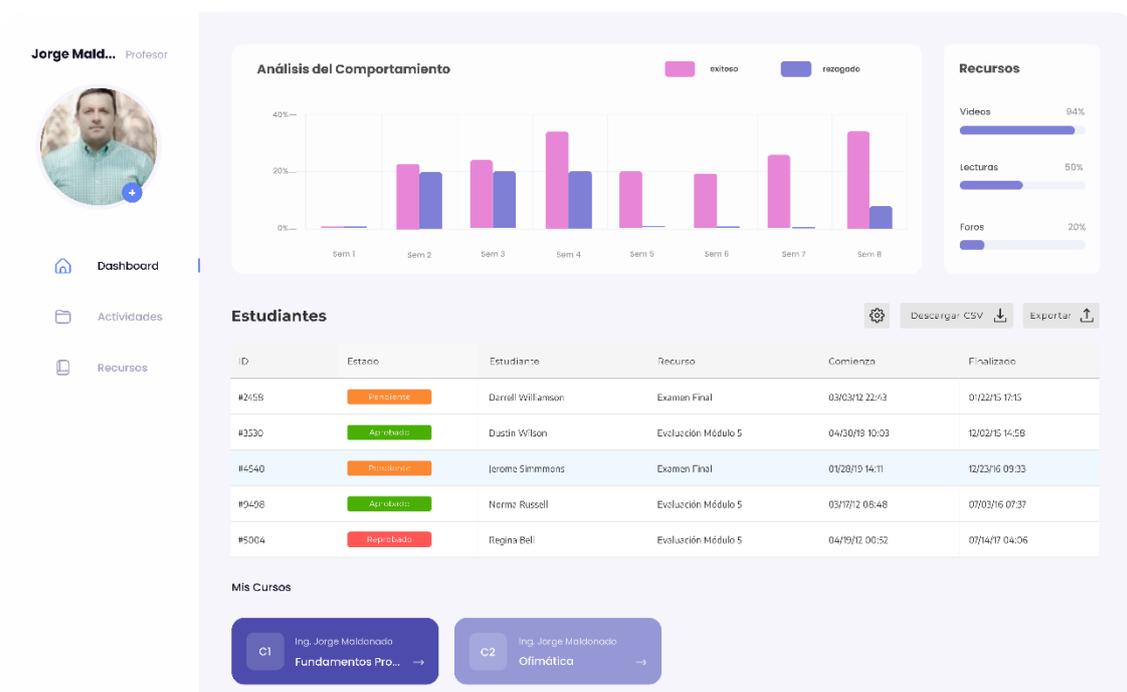


Figura 29. Prototipo del dashboard de la vista del docente.

5.5. Conclusiones

En este capítulo se aborda el tercer paso de la metodología de este trabajo de titulación, luego de haber realizado el análisis y selección de los indicadores obtenidos en el capítulo 4. Con el fin de seleccionar aquellas visualizaciones que cumplan con los objetivos y requerimientos expuestos en capítulos anteriores, se revisaron otras herramientas de LA, tales como Insights, el sistema de analíticas de datos de Open edX. Estas visualizaciones carecen de información detallada acerca del proceso de aprendizaje de cada estudiante, otras herramientas como ANALIZE y OXALIC se enfocan en el proceso de aprendizaje de los estudiantes, sin embargo, solo se enfocan en los maestros dejando a los estudiantes sin información acerca de su progreso en el curso.

De acuerdo a la guía para la elección de visualizaciones para dashboards de aprendizaje “Linking dashboard design and data visualization concepts”, nuestro componente de LA se deberá enfocar en los objetivos y necesidades de estudiantes y maestros, en las fuentes de datos, retroalimentación y autorregulación del aprendizaje y en la facilidad para medir el proceso de aprendizaje. Las visualizaciones contendrán información acerca de la cantidad de recursos completados por semanas, interacciones con los videos del curso, total de lecturas, problemas y videos completados a lo largo del curso, además

UCUENCA

del tiempo invertido. La mayoría de las visualizaciones ofrecen a estudiantes y maestros una perspectiva acerca del estado general del curso a través del promedio de los datos analizados por semanas.

Capítulo 6: Desarrollo del Xblock

6.1. Introducción

En este capítulo se ejecutan las etapas de implementación y evaluación del XBlock de la metodología usada en este trabajo de titulación. La primera sección de este capítulo muestra la arquitectura planteada para el desarrollo del XBlock y que tecnologías fueron empleadas para su desarrollo. Posteriormente, se ejecuta la etapa de evaluación de la metodología propuesta con el propósito de realizar una evaluación de las visualizaciones que son mostradas en el XBlock. Para llevar a cabo esta última etapa, se seleccionan a estudiantes de la Universidad de Cuenca para evaluar el componente desarrollado en este trabajo de titulación.

6.2. Implementación de las visualizaciones dentro de Open edX

En esta subsección se implementan las visualizaciones diseñadas en el capítulo anterior. Se revisa la arquitectura de los componentes de Open edX, llamados XBlock, así también, se presenta la arquitectura del componente desarrollado en este trabajo de titulación. Se detallan los pasos a seguir para la instalación y ejecución del componente en un ambiente de desarrollo. Finalmente, se describe el contexto en el que se evaluará las visualizaciones del componente de LA.

6.2.1. Arquitectura de un XBlock

Las tecnologías empleadas en el XBlock son las mismas que se usa en las aplicaciones web, donde se tienen múltiples componentes de manera independiente para ser mostrados en una sola página web (edX Inc, 2021). En el desarrollo no se necesita descargar toda la pila de ejecución de la plataforma edX ni conocer todas las tecnologías que emplea el XBlock, solo necesitan usar XBlock-sdk y realizar la implementación en cualquier versión de la plataforma que sea compatible con los XBlocks.

Las aplicaciones de tipo web que implementen XBlock API son consideradas XBlocks en entorno de ejecución (edX Inc, 2021). Estas aplicaciones web se pueden componer de páginas web que no necesitan comunicarse con los otros XBlocks que se están en tiempo de ejecución permitiendo integrar nuevas funcionalidades a la plataforma. Estas aplicaciones web en miniatura cuentan con una capa de almacenamiento para mantener su estado, controladores para procesar las acciones del usuario y además cuentan con vistas para poder ser renderizadas (edX Inc, 2021). Estas aplicaciones web en miniatura pueden

UCUENCA

representar una pequeña parte de una página completa como párrafos de un texto, un video, campos de opción múltiple o estas pueden ser tan grandes como un capítulo entero o inclusive el curso completo (edX Inc, 2021).

En la primera de tres capas de la arquitectura propuesta para el XBlock de nuestro trabajo de titulación denominado XLEA (XBlock for LEarning Analytics), tenemos la parte visual. En esta primera capa es donde los estudiantes y los profesores interactúan con las visualizaciones, esta primera capa está compuesta por HTML, CSS y Javascript, los tres pilares que forman una aplicación web básica, esta aplicación integra XBlock API para poder acceder a la configuración con la que cuenta la plataforma de Open edX.

Descendiendo por la siguiente capa propuestas tenemos la comunicación con el código base de la plataforma de Open edX por medio del XBlock Runtime. En esta capa es donde el LMS y CMS requiere que los XBlocks tengan las siguientes propiedades, la primera es tener un método de la vista denominado *student_view* que en el caso del LMS esta se pueda interactuar, mientras que en el CMS el instructor pueda configurar, además deben de tener una propiedad llamada *has_score* con un valor de True si el XBlock permite realizar calificaciones, caso contrario este valor deberá ser configurado como False.

En la última capa de la arquitectura propuesta es la comunicación con los sistemas de persistencia de base de datos para recuperar la información almacenada. La comunicación con las bases de datos tanto de MySQL como de MongoDB se lo realiza por medio del XBlock Runtime que permite acceder en tiempo de ejecución a los datos almacenados y poder comunicar con las capas superiores. En esta capa es donde el XBlock puede acceder a la información procesada de la interacción de los estudiantes, pues previamente se ha almacenado en la base de datos de MySQL, empleando el script desarrollado, que los adjunta en el **Anexo 5**. La representación gráfica de la arquitectura propuesta con todas las capas antes mencionadas se muestra en la **Figura 30**.

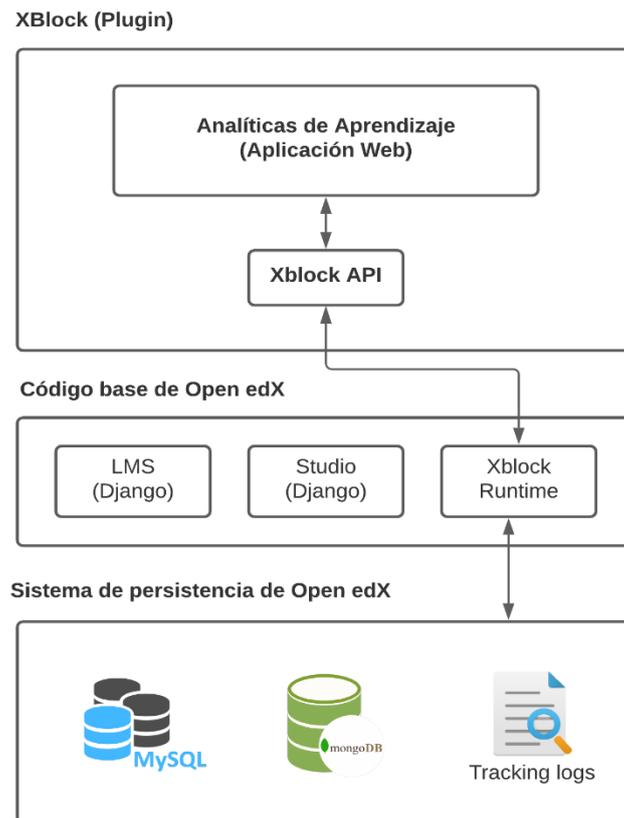


Figura 30. Arquitectura propuesta para el XBlock de LA.

6.2.2. Funcionalidades

Empleando la arquitectura antes mencionada, en el XBlock a desarrollar se incorporan funcionalidades adicionales a las ya existentes dentro de la plataforma Open edX. De esta manera tanto los estudiantes y maestros pueden tomar decisiones informadas en base a la información mostrada en el dashboard.

Las funcionalidades que brinda el componente de LA desarrollado son:

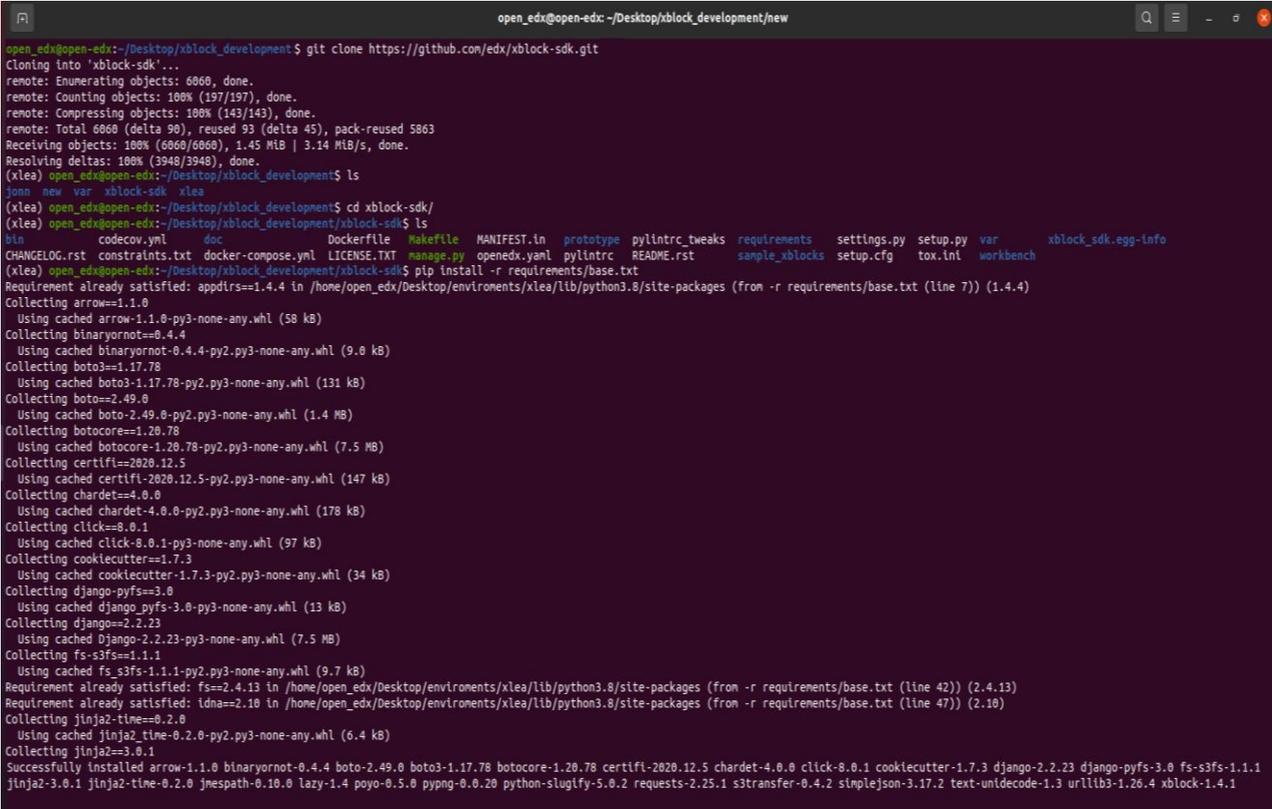
- Vista general de las interacciones de los estudiantes con los recursos que se encuentran dentro del curso.
- Medir y comparar los logros conseguidos de un estudiante con la media del curso.
- Comparación del tiempo dedicado por un estudiante con el tiempo promedio que ocupan los estudiantes en una semana de actividades en la plataforma.

6.2.3. Desarrollo del XBlock

UCUENCA

Un kit de desarrollo de software (SDK por sus abreviaturas en inglés de Software Development Kit) es un paquete que agrupa herramientas y datos que facilita a los desarrolladores la elaboración de programas en un lenguaje concreto o para una plataforma en específico (IONOS, 2019). La plataforma Open edX cuenta con un SDK para el desarrollo de los XBlocks denominada XBlock SDK, en este SDK se tiene tres componentes principales, los mismos que son: a) Una herramienta de creación de una estructura básica de un XBlock; b) Un entorno de ejecución para probar los XBlocks en desarrollo; c) Ejemplos ilustrativos de la creación de los XBlocks como un punto de partida para nuevos desarrollos (Open edX, 2022a).

El XBlock SDK no viene instalado por defecto cuando se instala la plataforma Open edX, por lo que se debe de realizar pasos adicionales para su instalación. El primer paso para su instalación es crear un entorno virtual de Python y activarlo; el segundo paso es clonar el repositorio de github donde se encuentra el código del SDK desde el siguiente enlace: <https://github.com/openedx/xblock-sdk> e instalar todos los requerimientos necesarios para su ejecución (Open edX, 2022a). La realización de este proceso de instalación se lo muestra en la **Figura 31**.



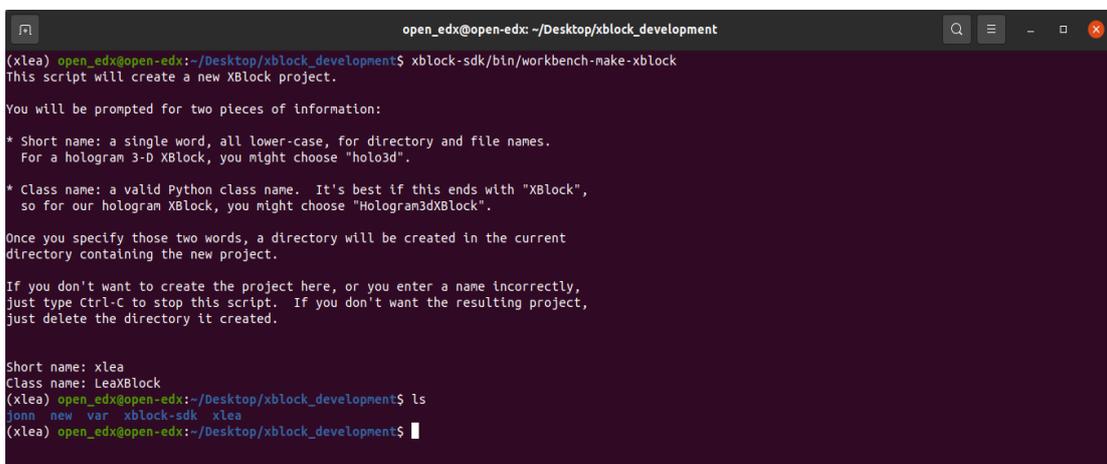
```
open_edx@open-edx: ~/Desktop/xblock_development/new
open_edx@open-edx:~/Desktop/xblock_development$ git clone https://github.com/openedx/xblock-sdk.git
Cloning into 'xblock-sdk'...
remote: Enumerating objects: 6060, done.
remote: Counting objects: 100% (197/197), done.
remote: Compressing objects: 100% (143/143), done.
remote: Total 6060 (delta 90), reused 93 (delta 45), pack-reused 5863
Receiving objects: 100% (6060/6060), 1.45 MiB | 3.14 MiB/s, done.
Resolving deltas: 100% (3948/3948), done.
(xlea) open_edx@open-edx:~/Desktop/xblock_development$ ls
jonn new var xblock-sdk xlea
(xlea) open_edx@open-edx:~/Desktop/xblock_development$ cd xblock-sdk/
(xlea) open_edx@open-edx:~/Desktop/xblock_development/xblock-sdk$ ls
bin          codecov.yml  doc          Dockerfile  Makefile    MANIFEST.in  prototype  pylintrc  tweaks  requirements  settings.py  setup.py  var          xblock_sdk.egg-info
CHANGELOG.rst  constraints.txt  docker-compose.yml  LICENSE.TXT  manage.py  openedx.yaml  pylintrc  README.rst  sample_xblocks  setup.cfg  tox.ini  workbench
(xlea) open_edx@open-edx:~/Desktop/xblock_development/xblock-sdk$ pip install -r requirements/base.txt
Requirement already satisfied: appdirs==1.4.4 in /home/open_edx/Desktop/enviroments/xlea/lib/python3.8/site-packages (from -r requirements/base.txt (line 7)) (1.4.4)
Collecting arrow==1.1.0
  Using cached arrow-1.1.0-py3-none-any.whl (58 kB)
Collecting binaryornot==0.4.4
  Using cached binaryornot-0.4.4-py2.py3-none-any.whl (9.0 kB)
Collecting boto3==1.17.78
  Using cached boto3-1.17.78-py2.py3-none-any.whl (131 kB)
Collecting boto==2.49.0
  Using cached boto-2.49.0-py2.py3-none-any.whl (1.4 MB)
Collecting botocore==1.20.78
  Using cached botocore-1.20.78-py2.py3-none-any.whl (7.5 MB)
Collecting certifi==2020.12.5
  Using cached certifi-2020.12.5-py2.py3-none-any.whl (147 kB)
Collecting chardet==4.0.0
  Using cached chardet-4.0.0-py2.py3-none-any.whl (178 kB)
Collecting click==8.0.1
  Using cached click-8.0.1-py3-none-any.whl (97 kB)
Collecting cookiecutter==1.7.3
  Using cached cookiecutter-1.7.3-py2.py3-none-any.whl (34 kB)
Collecting django-pyfs==3.0
  Using cached django_pyfs-3.0-py3-none-any.whl (13 kB)
Collecting django==2.2.23
  Using cached Django-2.2.23-py3-none-any.whl (7.5 MB)
Collecting fs-s3fs==1.1.1
  Using cached fs_s3fs-1.1.1-py2.py3-none-any.whl (9.7 kB)
Requirement already satisfied: fs==2.4.13 in /home/open_edx/Desktop/enviroments/xlea/lib/python3.8/site-packages (from -r requirements/base.txt (line 42)) (2.4.13)
Requirement already satisfied: idna==2.10 in /home/open_edx/Desktop/enviroments/xlea/lib/python3.8/site-packages (from -r requirements/base.txt (line 47)) (2.10)
Collecting jinja2-ttne==0.2.0
  Using cached jinja2_ttne-0.2.0-py2.py3-none-any.whl (6.4 kB)
Collecting jinja2==3.0.1
  Using cached Jinja2-3.0.1-py3-none-any.whl (133 kB)
Successfully installed arrow-1.1.0 binaryornot-0.4.4 boto-2.49.0 boto3-1.17.78 botocore-1.20.78 certifi-2020.12.5 chardet-4.0.0 click-8.0.1 cookiecutter-1.7.3 django-2.2.23 django-pyfs-3.0 fs-s3fs-1.1.1
jinja2-3.0.1 jinja2-ttne-0.2.0 jmespath-0.10.0 lazy-1.4 poyo-0.5.0 pyng-0.8.20 python-slugify-5.0.2 requests-2.25.1 s3transfer-0.4.2 simplejson-3.17.2 text-unidecode-1.3 urllib3-1.26.4 xblock-1.4.1
```

Figura 31. Instalación del XBlock SDK.

UCUENCA

La creación de un XBlock se lo realiza empleando XBlock SDK, el mismo que crea una estructura definida por los desarrolladores de la plataforma Open edX. En la estructura básica se crean las carpetas y archivos necesarios, pero para poder obtener esta estructura primero se debe de activar el entorno virtual (con los requerimientos previamente instalados) y ejecutar el comando que se muestra en la **Figura 32**, después de ejecutar dicho comando tendremos una interacción con la terminal donde se ingresara información en cuanto al nombre del XBlock y el nombre de la clase para nuestro XBlock. Después de los pasos previamente mencionados contaremos con la estructura básica de nuestro XBlock, tal como se lo muestra en la **Figura 33**.

Dentro de esta estructura se encuentra el archivo de configuración del XBlock denominado `setup.py` y una carpeta con el mismo nombre asignado, dentro de esa carpeta tenemos la carpeta `static` donde se encuentran las tecnologías básicas de un proyecto web las cuales son *JavaScript*, *CSS* y *HTML*; también se encuentra otra carpeta `translations` la cual es usada para extraer las cadenas de texto para traducir al lenguaje respectivo y ofrecer las internacionalización; además tenemos el archivo `xlea.py` en el cual se define los *campos*, *vistas*, *manejadores* y el escenario de trabajo en general.



```
open_edx@open-edx: ~/Desktop/xblock_development
(xlea) open_edx@open-edx:~/Desktop/xblock_development$ xblock-sdk/bin/workbench-make-xblock
This script will create a new XBlock project.

You will be prompted for two pieces of information:

* Short name: a single word, all lower-case, for directory and file names.
  For a hologram 3-D XBlock, you might choose "holo3d".

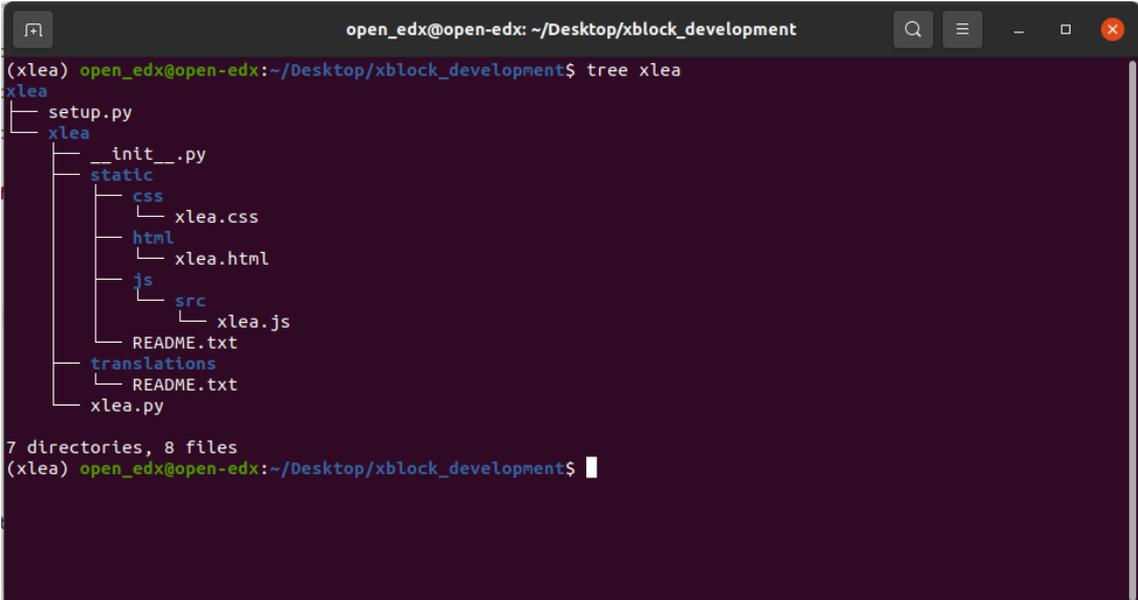
* Class name: a valid Python class name. It's best if this ends with "XBlock",
  so for our hologram XBlock, you might choose "Hologram3dXBlock".

Once you specify those two words, a directory will be created in the current
directory containing the new project.

If you don't want to create the project here, or you enter a name incorrectly,
just type Ctrl-C to stop this script. If you don't want the resulting project,
just delete the directory it created.

Short name: xlea
Class name: LeaXBlock
(xlea) open_edx@open-edx:~/Desktop/xblock_development$ ls
jonn new var xblock-sdk xlea
(xlea) open_edx@open-edx:~/Desktop/xblock_development$
```

Figura 32. Comando para crear un XBlock usando el XBlock SDK.



```
open_edx@open-edx: ~/Desktop/xblock_development
(xlea) open_edx@open-edx:~/Desktop/xblock_development$ tree xlea
xlea
├── setup.py
├── xlea
│   ├── __init__.py
│   ├── static
│   │   ├── css
│   │   │   └── xlea.css
│   │   ├── html
│   │   │   └── xlea.html
│   │   └── js
│   │       └── src
│   │           └── xlea.js
│   ├── README.txt
│   ├── translations
│   │   └── README.txt
│   └── xlea.py
└── 7 directories, 8 files
(xlea) open_edx@open-edx:~/Desktop/xblock_development$
```

Figura 33. Estructura de las carpeta y archivos de un XBlock.

Para la instalación y ejecución del XBlock desarrollado se emplea el SDK server propio del XBlock SDK. En el mismo directorio creado para el desarrollo se debe de instalar el XBlock usando los comandos que se muestran en la **Figura 34**, estos comandos permiten instalar el XBlock en el entorno de pruebas y visualizarlos por medio de una interfaz gráfica cuando el SDK server se encuentra en ejecución. Al ingresar en la interfaz gráfica principal nos pedirá que ingresemos el nombre del XBlock que deseamos visualizar, en nuestro caso es XBlock for Learning Analytics (XLEA, por sus siglas en inglés). Después de dar clic nos mostrará el dashboard con los indicadores de LA.

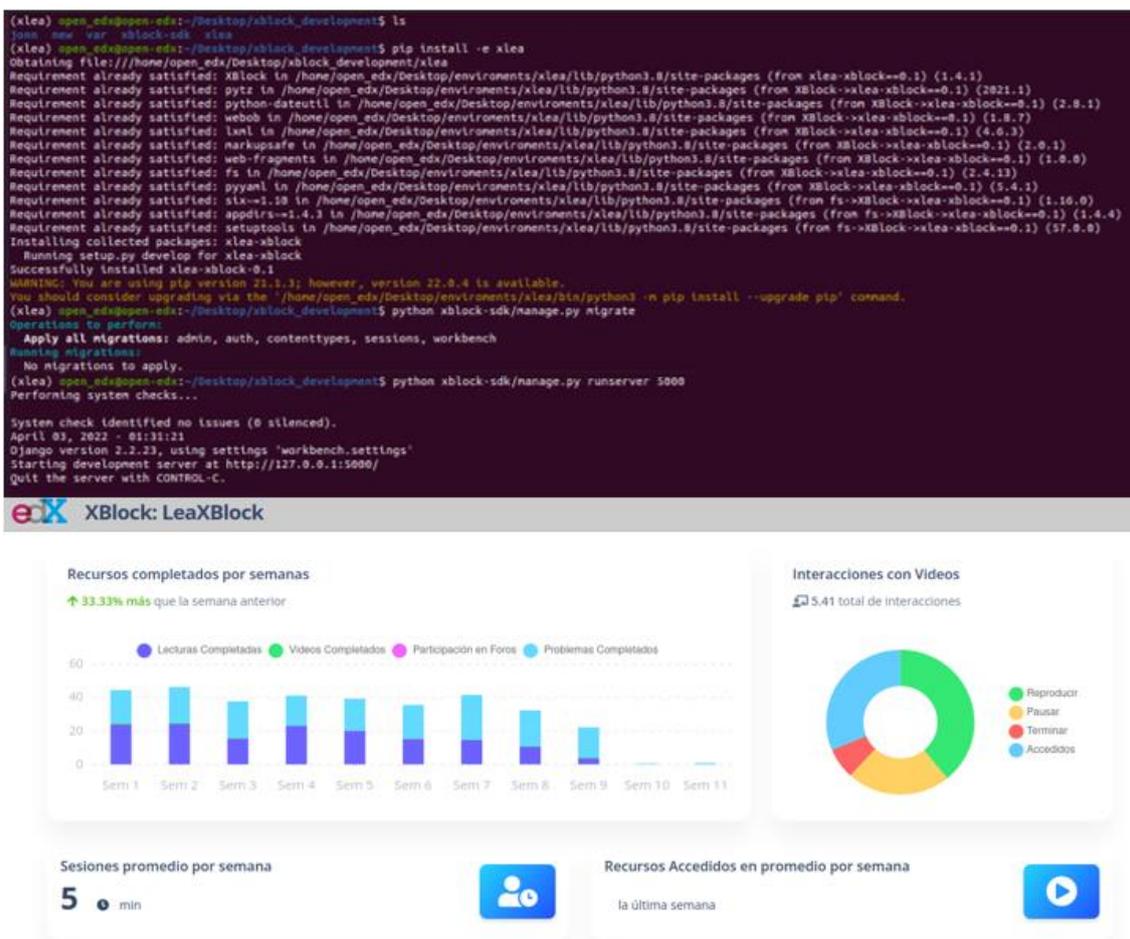


Figura 34. Instalación y ejecución del XBlock dentro del SDK Server.

6.2.4. Visualizaciones del XBlock desarrollado en este trabajo de titulación

En esta subsección se presentan las visualizaciones que se implementarán en el dashboard de LA antes mencionado enfocadas en dos tipos de vista, la del estudiante y la del docente. Las visualizaciones se diseñaron en base a los objetivos y requerimientos de este trabajo de titulación; así también se incorporaron los indicadores obtenidos del capítulo 4 y los prototipos diseñados en el capítulo 5.

En la vista del estudiante, se muestra una gráfica de línea (ver **Figura 35**) que representa el rendimiento general del estudiante por semanas. Los indicadores que se utilizaron para el diseño de este gráfico fueron la cantidad de recursos a los que ha accedido el estudiante a lo largo de una semana, es decir, lecturas, videos, foros y exámenes.



Figura 35. Rendimiento general por semanas. Vista del estudiante.

Adicionalmente se compara el rendimiento individual del estudiante contra el rendimiento del curso en promedio, con el fin de dar a los estudiantes información acerca del estado de su proceso de aprendizaje. En la **Figura 36** se muestra el tiempo invertido del estudiante en un curso por semanas. Para el diseño de esta visualización se incorporaron indicadores de la cantidad de sesiones que realiza un estudiante por semana y se comparó con la media del curso.

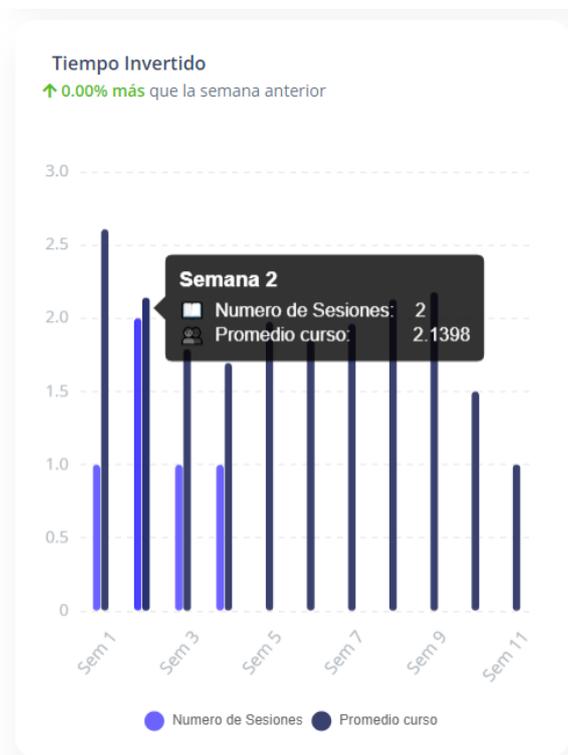


Figura 36. Tiempo invertido en el curso por semanas. Vista del estudiante.

UCUENCA

En la **Figura 37** se muestran indicadores relacionados a la cantidad de recursos completados por el estudiante a lo largo del curso. Los recursos a considerar son las lecturas, videos y problemas. Cada uno visualiza la cantidad de interacciones con estos recursos, es decir, cuando un estudiante inicia, finaliza y repite un recurso en específico. De igual manera, se realiza una comparación con la media del curso.



Figura 37. Recursos completados a lo largo del curso. Vista del estudiante.

Para la vista del docente, las visualizaciones se diseñaron en base al promedio de las interacciones de todos los estudiantes del curso. En la **Figura 38** se muestran los recursos completados por semanas de todos los estudiantes del curso, separados en lecturas, videos, foros y problemas.



Figura 38. Recursos completados por semanas de todos los estudiantes del curso. Vista del docente.

En la **Figura 39** se muestra una gráfica de dona de las interacciones con videos de los estudiantes a lo largo del curso. Las interacciones se agrupan en cantidad de reproducciones, pausas, finalización y acceso de videos. También se visualiza el número total de visualizaciones.



Figura 39. Interacciones con videos de todos los estudiantes del curso. Vista del docente.

UCUENCA

En la **Figura 40** se visualizan las sesiones en promedio por semana y su duración, así también la cantidad de recursos accedidos en promedio por semana de todos los estudiantes del curso. Además, se muestra la duración en promedio de las sesiones de los estudiantes.



Figura 40. Sesiones promedio por semana y recursos accedidos en promedio por semana. Vista del docente.

Finalmente, en la **Figura 41** se muestra una tabla con las estadísticas individuales de cada estudiante del curso. Se incluyen indicadores como el número de recursos accedidos, lecturas, videos, problemas, foros, número de sesiones y el tiempo invertido en el curso.

Estadísticas del curso
 Total de estudiantes 273

ESTUDIANTE	RECURSOS ACCEDIDOS	LECTURAS	VIDEOS	PROBLEMAS	FOROS	SESIONES
jorge.maldonado@ucuenca.edu.ec	15.75	10.12	0.25	0.75	0.00	3.38
LORENA CATALINASIGUENZA GUZMAN lorena.siguenza@ucuenca.edu.ec	14.00	6.00	0.00	0.00	0.00	5.00
JORGEMALDONADO jjmm1280@gmail.com	11.00	6.00	0.00	0.00	0.00	1.00
JUAN LUIS ALFREDOESCALANTE CALDERON juan.escalante@ucuenca.edu.ec	62.00	21.86	0.57	20.29	0.00	2.57
darwin.chavez0601@ucuenca.edu.ec	38.88	16.12	0.50	17.25	0.00	2.62

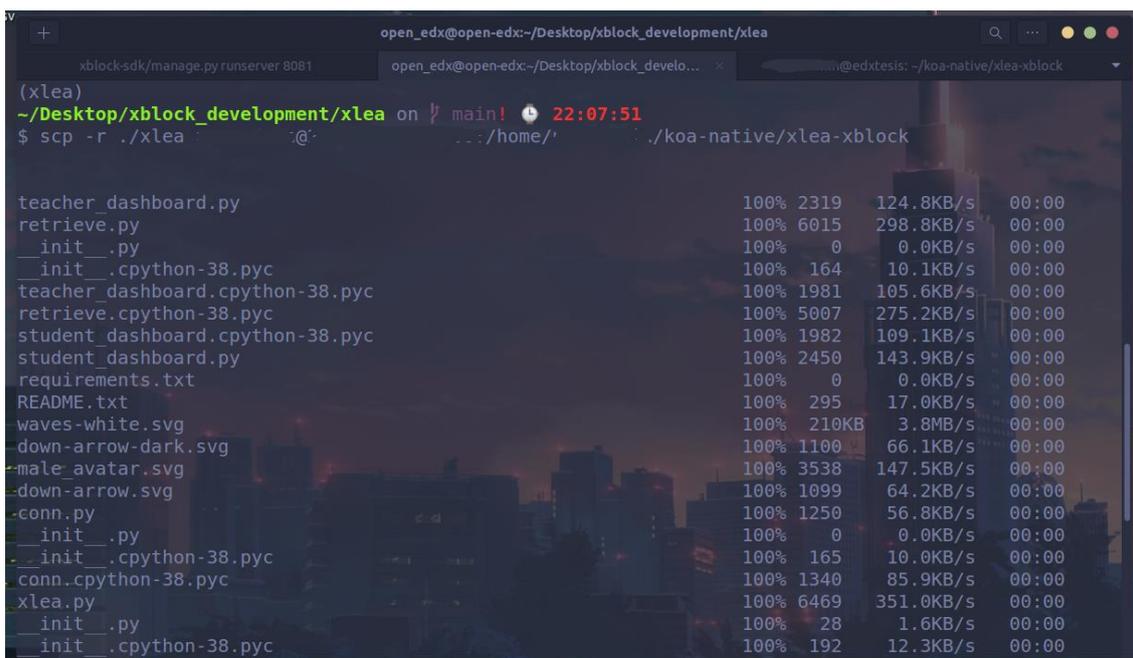
Figura 41. Estadísticas de los indicadores de cada estudiante del curso. Vista del docente.

6.3. Instalación del XBlock en un ambiente de producción

En esta sección se introduce al lector los pasos que se siguieron en este trabajo de titulación para desplegar el XBlock en un servidor de producción. Este servidor fue provisto por la Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia (CEDIA), el mismo que será utilizado para evaluar el dashboard de LA. Se puede acceder al servidor de producción en la siguiente URL: <https://test-xlearning.cedia.edu.ec>. Los datos de los cursos para alimentar el algoritmo de analíticas fueron dispuestos por la Universidad de Cuenca. En este contexto se procedió a instalar el componente XBlock en un servidor Ubuntu 20.04 con la versión Koa.3 de Open edX.

Es importante mencionar que para el correcto funcionamiento del componente de LA se deberá configurar el script de análisis de datos desarrollado en este trabajo de titulación. Este script debe ser colocado en el servidor de producción con permisos de root. Esto es necesario debido a que el script se ejecuta en periodos semanales para procesar los trackings logs y luego almacenarlos en la base de datos de Open edX. Los detalles de cómo se trabajaron estos datos se explican en el capítulo 4, así también se pone a disposición del lector el código desarrollado por los autores en el **Anexo 5**.

Para la correcta instalación del servidor Open edX se recomienda seguir los pasos descritos en la documentación oficial (Blackwell, 2021). Luego de instalar la versión Koa.3, que se utilizó para este trabajo de titulación, se procede a copiar (ver **Figura 42**) los archivos del proyecto XLEA que contiene nuestro componente de LA. Los pasos descritos a continuación se tomaron de los foros oficiales de la plataforma Open edX (*XBlock Installation in Open EdX Lilac - Open EdX Discussions*, n.d.).

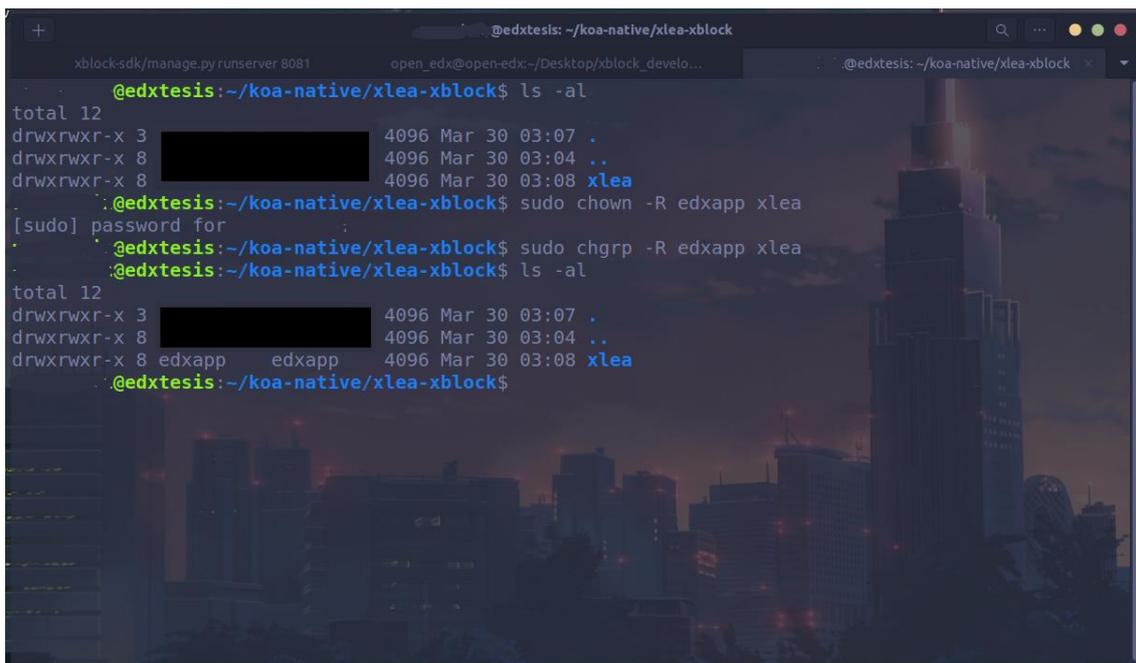


```
open_edx@open-edx:~/Desktop/xblock_development/xlea
(xlea)
~/Desktop/xblock_development/xlea on main! 22:07:51
$ scp -r ./xlea @: /home/ /koa-native/xlea-xblock

teacher_dashboard.py          100% 2319   124.8KB/s   00:00
retrieve.py                   100% 6015   298.8KB/s   00:00
__init__.py                   100% 0      0.0KB/s     00:00
__init__.cpython-38.pyc      100% 164    10.1KB/s    00:00
teacher_dashboard.cpython-38.pyc 100% 1981   105.6KB/s   00:00
retrieve.cpython-38.pyc      100% 5007   275.2KB/s   00:00
student_dashboard.cpython-38.pyc 100% 1982   109.1KB/s   00:00
student_dashboard.py         100% 2450   143.9KB/s   00:00
requirements.txt             100% 0      0.0KB/s     00:00
README.txt                    100% 295    17.0KB/s    00:00
waves-white.svg              100% 210KB   3.8MB/s     00:00
down-arrow-dark.svg          100% 1100    66.1KB/s    00:00
male_avatar.svg              100% 3538   147.5KB/s   00:00
down-arrow.svg               100% 1099    64.2KB/s    00:00
conn.py                       100% 1250    56.8KB/s    00:00
__init__.py                   100% 0      0.0KB/s     00:00
__init__.cpython-38.pyc      100% 165    10.0KB/s    00:00
conn.cpython-38.pyc          100% 1340    85.9KB/s    00:00
xlea.py                       100% 6469   351.0KB/s   00:00
__init__.py                   100% 28     1.6KB/s     00:00
__init__.cpython-38.pyc      100% 192    12.3KB/s    00:00
```

Figura 42. Migración del código de XLEA al servidor de producción.

Para acceder al servidor de producción se puso a nuestra disposición las respectivas llaves de acceso por medio de SSH. Una vez el código del XBlock ha sido migrado al servidor de producción se procede a cambiar los permisos del directorio raíz del proyecto (ver **Figura 43**). Se debe configurar como nuevo dueño y grupo al usuario “edxapp”, este usuario que se crea por defecto al momento de instalar Open edX.



```
@edxtesis: ~/koa-native/xlea-xblock
xblock-sdk/manage.py runserver 8081  open_edx@open-edx:~/Desktop/xblock_develo...  @edxtesis: ~/koa-native/xlea-xblock
@edxtesis:~/koa-native/xlea-xblock$ ls -al
total 12
drwxrwxr-x 3  4096 Mar 30 03:07 .
drwxrwxr-x 8  4096 Mar 30 03:04 ..
drwxrwxr-x 8  4096 Mar 30 03:08 xlea
@edxtesis:~/koa-native/xlea-xblock$ sudo chown -R edxapp xlea
[sudo] password for :
@edxtesis:~/koa-native/xlea-xblock$ sudo chgrp -R edxapp xlea
@edxtesis:~/koa-native/xlea-xblock$ ls -al
total 12
drwxrwxr-x 3  4096 Mar 30 03:07 .
drwxrwxr-x 8  4096 Mar 30 03:04 ..
drwxrwxr-x 8 edxapp  edxapp  4096 Mar 30 03:08 xlea
@edxtesis:~/koa-native/xlea-xblock$
```

Figura 43. Otorgar permisos al usuario “edxapp” sobre el directorio del componente de LA.

Luego de otorgar los respectivos permisos al directorio del componente XBlock, se procede a instalar las dependencias necesarias para el correcto funcionamiento del mismo. Se debe ingresar a la consola bash del usuario “edxapp” como se muestra en la **Figura 44**. Una vez iniciada la sesión en la consola, el usuario deberá dirigirse al directorio /edx/app/edxapp para activar el ambiente de desarrollo propio de Open edX. Luego se podrá instalar las dependencias con ayuda de un archivo de configuración llamado requirements.txt que viene incluido en el directorio del XBlock desarrollado.

```
@edxtesis: ~/koa-native/xlea-xblock
xblock-sdk/manage.py runcserver 8081
open_edx@open-edx:~/Desktop/xblock_develo...
@edxtesis: ~/koa-native/xlea-xblock

@edxtesis:~/koa-native/xlea-xblock$ sudo su edxapp -s /bin/bash
edxapp@edxtesis:/home/.../koa-native/xlea-xblock$ cd /edx/app/edxapp/
edxapp@edxtesis:~$ source edxapp_env
edxapp@edxtesis:~$ pip install -r /home/.../koa-native/xlea-xblock/xlea/requirements.txt
Requirement already satisfied: appdirs==1.4.4 in ./venvs/edxapp/lib/python3.8/site-packages (from -
r /home/.../koa-native/xlea-xblock/xlea/requirements.txt (line 2)) (1.4.4)
Collecting boto3==1.20.37
  Downloading boto3-1.20.37-py3-none-any.whl (131 kB)
  |-----| 131 kB 735 kB/s
Collecting botocore==1.23.37
  Downloading botocore-1.23.37-py3-none-any.whl (8.5 MB)
  |-----| 8.5 MB 4.1 MB/s
Collecting certifi==2021.5.30
  Downloading certifi-2021.5.30-py2.py3-none-any.whl (145 kB)
  |-----| 145 kB 94.7 MB/s
Requirement already satisfied: chardet==3.0.4 in ./venvs/edxapp/lib/python3.8/site-packages (from -
r /home/.../koa-native/xlea-xblock/xlea/requirements.txt (line 6)) (3.0.4)
Collecting Django==2.2.14
  Downloading Django-2.2.14-py3-none-any.whl (7.5 MB)
  |-----| 7.5 MB 38.0 MB/s
Collecting django-pyfs==2.0
  Downloading django_pyfs-2.0-py3-none-any.whl (11 kB)
Collecting fs==2.4.13
  Downloading fs-2.4.13-py2.py3-none-any.whl (131 kB)
  |-----| 131 kB 36.1 MB/s
Collecting fs-s3fs==1.1.1
```

Figura 44. Instalación de las dependencias del componente de LA.

Cuando todas las dependencias han sido instaladas satisfactoriamente, se procede a instalar el XBlock dentro de Open edX (ver **Figura 45**).

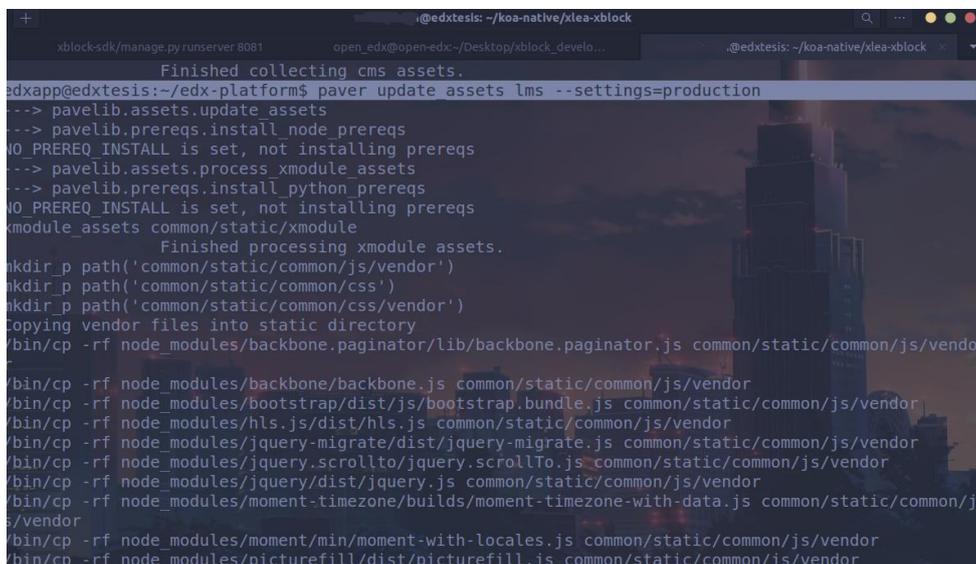
```
@edxtesis: ~/koa-native/xlea-xblock
xblock-sdk/manage.py runcserver 8081
open_edx@open-edx:~/Desktop/xblock_develo...
@edxtesis: ~/koa-native/xlea-xblock

edxapp@edxtesis:~$ pip install /home/.../koa-native/xlea-xblock/xlea
Processing /home/.../koa-native/xlea-xblock/xlea
Requirement already satisfied: XBlock in ./venvs/edxapp/lib/python3.8/site-packages (from xlea-xblo
ck==0.11) (1.4.2)
Requirement already satisfied: python-dateutil in ./venvs/edxapp/lib/python3.8/site-packages (from
XBlock->xlea-xblock==0.11) (2.8.1)
Requirement already satisfied: pyyaml in ./venvs/edxapp/lib/python3.8/site-packages (from XBlock->x
lea-xblock==0.11) (5.4.1)
Requirement already satisfied: lxml in ./venvs/edxapp/lib/python3.8/site-packages (from XBlock->xle
a-xblock==0.11) (4.6.3)
Requirement already satisfied: web-fragments in ./venvs/edxapp/lib/python3.8/site-packages (from XB
lock->xlea-xblock==0.11) (1.0.0)
Requirement already satisfied: fs in ./venvs/edxapp/lib/python3.8/site-packages (from XBlock->xlea
-xblock==0.11) (2.4.13)
Requirement already satisfied: markupsafe in ./venvs/edxapp/lib/python3.8/site-packages (from XBloc
k->xlea-xblock==0.11) (2.0.1)
Requirement already satisfied: webob in ./venvs/edxapp/lib/python3.8/site-packages (from XBlock->xl
ea-xblock==0.11) (1.8.7)
Requirement already satisfied: pytz in ./venvs/edxapp/lib/python3.8/site-packages (from XBlock->xle
a-xblock==0.11) (2021.1)
Requirement already satisfied: six>=1.5 in ./venvs/edxapp/lib/python3.8/site-packages (from python-
dateutil->XBlock->xlea-xblock==0.11) (1.16.0)
Requirement already satisfied: appdirs==1.4.3 in ./venvs/edxapp/lib/python3.8/site-packages (from f
s->XBlock->xlea-xblock==0.11) (1.4.4)
Requirement already satisfied: setuptools in ./venvs/edxapp/lib/python3.8/site-packages (from fs->X
Block->xlea-xblock==0.11) (60.9.3)
```

Figura 45. Instalación del componente de LA en Open edX.

UCUENCA

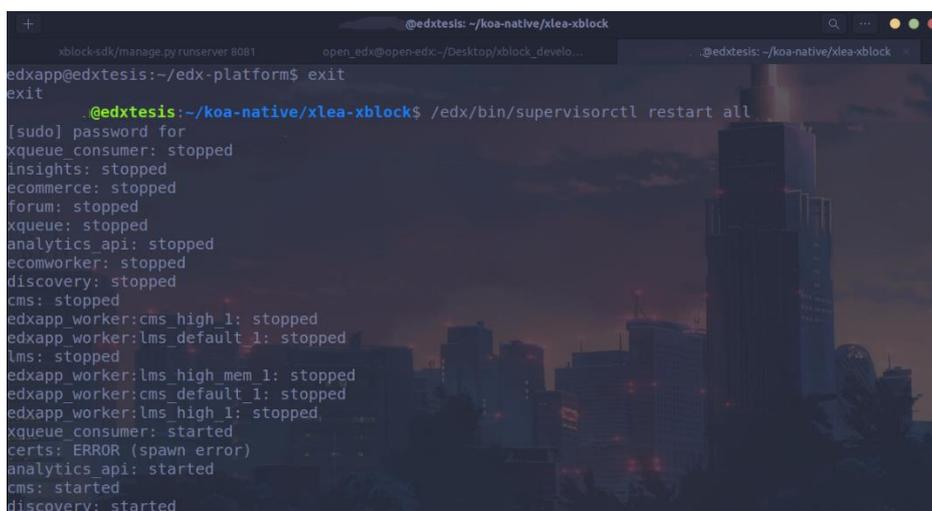
Como último paso, el usuario debe actualizar los recursos y archivos estáticos del servidor Open edX. Para esto, se debe ejecutar el comando “paver” el cuál se encarga de gestionar el proceso de actualización, este comando debe ser ejecutado dos veces, para actualizar los archivos en el LMS (ver **Figura 46**) y para actualizar el CMS (se debe cambiar el parámetro “lms” por “cms”).



```
Finished collecting cms assets.
edxapp@edxtestis:~/edx-platform$ paver update assets lms --settings-production
--> pavelib.assets.update_assets
--> pavelib.prereqs.install_node_prereqs
NO_PREREQ_INSTALL is set, not installing prereqs
--> pavelib.assets.process_xmodule_assets
--> pavelib.prereqs.install_python_prereqs
NO_PREREQ_INSTALL is set, not installing prereqs
xmodule_assets common/static/xmodule
Finished processing xmodule assets.
mkdir_p path('common/static/common/js/vendor')
mkdir_p path('common/static/common/css')
mkdir_p path('common/static/common/css/vendor')
Copying vendor files into static directory
/bin/cp -rf node_modules/backbone/paginator/lib/backbone.paginator.js common/static/common/js/vendor
/bin/cp -rf node_modules/backbone/backbone.js common/static/common/js/vendor
/bin/cp -rf node_modules/bootstrap/dist/js/bootstrap.bundle.js common/static/common/js/vendor
/bin/cp -rf node_modules/hls.js/dist/hls.js common/static/common/js/vendor
/bin/cp -rf node_modules/jquery-migrate/dist/jquery-migrate.js common/static/common/js/vendor
/bin/cp -rf node_modules/jquery.scrollTo/jquery.scrollTo.js common/static/common/js/vendor
/bin/cp -rf node_modules/jquery/dist/jquery.js common/static/common/js/vendor
/bin/cp -rf node_modules/moment-timezone/builds/moment-timezone-with-data.js common/static/common/js/vendor
/bin/cp -rf node_modules/moment/min/moment-with-locales.js common/static/common/js/vendor
/bin/cp -rf node_modules/picturefill/dist/picturefill.js common/static/common/js/vendor
```

Figura 46. Actualización del LMS en Open edX.

Finalmente, una vez actualizados el LMS y CMS de Open edX, se deben reiniciar todos los servicios para que los cambios se desplieguen a producción. El comando para reiniciar todos los servicios se muestra en la **Figura 47**.



```
exit
exit
@edxtestis:~/koa-native/xlea-xblock$ /edx/bin/supervisorctl restart all
[sudo] password for
xqueue_consumer: stopped
insights: stopped
ecommerce: stopped
forum: stopped
xqueue: stopped
analytics_api: stopped
ecomworker: stopped
discovery: stopped
cms: stopped
edxapp_worker:cms_high_1: stopped
edxapp_worker:lms_default_1: stopped
lms: stopped
edxapp_worker:lms_high_mem_1: stopped
edxapp_worker:cms_default_1: stopped
edxapp_worker:lms_high_1: stopped
xqueue_consumer: started
certs: ERROR (spawn error)
analytics_api: started
cms: started
discovery: started
```

Figura 47. Reinicio de todos los servicios de Open edX.

6.4. Ejecución del XBlock en un ambiente de producción

En esta sección se detallan los pasos a seguir para habilitar el dashboard de LA en la interfaz del LMS de Open edX. Una vez habilitado el XBlock, estudiantes y maestros tendrán acceso a las visualizaciones desarrolladas en este trabajo de titulación. Para habilitar el XBlock se siguieron los pasos descritos en la documentación oficial de Open edX (*Deploy Your XBlock in Devstack — Open EdX XBlock Tutorial Documentation*, n.d.).

Primeramente, el usuario deberá iniciar sesión como administrador en el dashboard del LMS de Open edX (ver **Figura 48**).

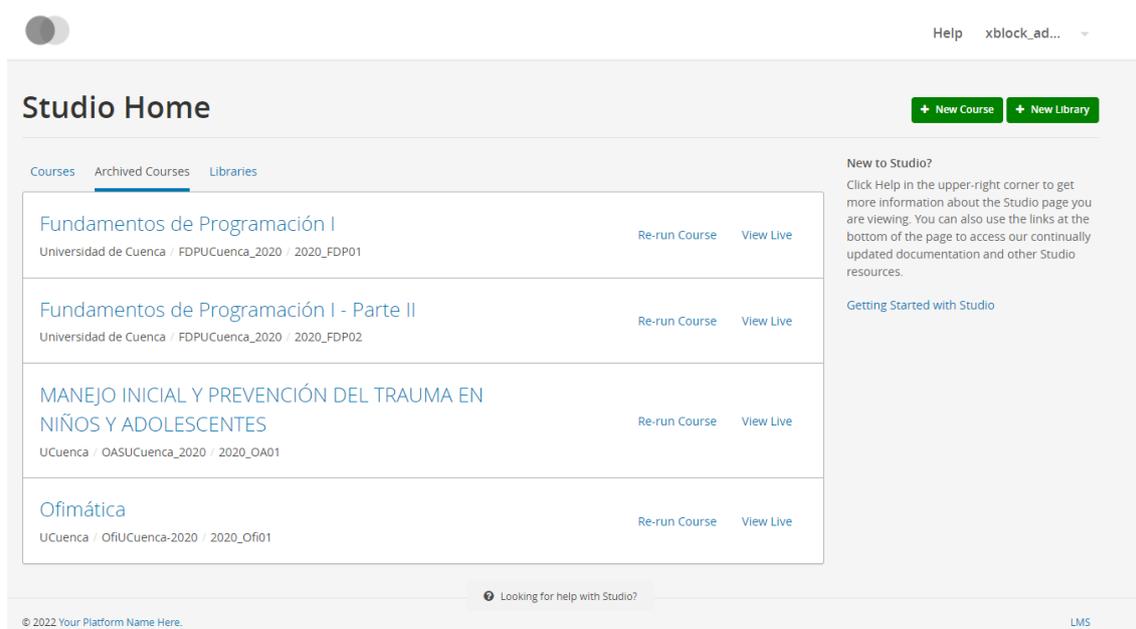


Figura 48. Inicio de sesión en el dashboard LMS del Open edX.

Luego hay que seleccionar el curso en el que se desea habilitar el XBlock, tal como se muestra en la **Figura 49**.

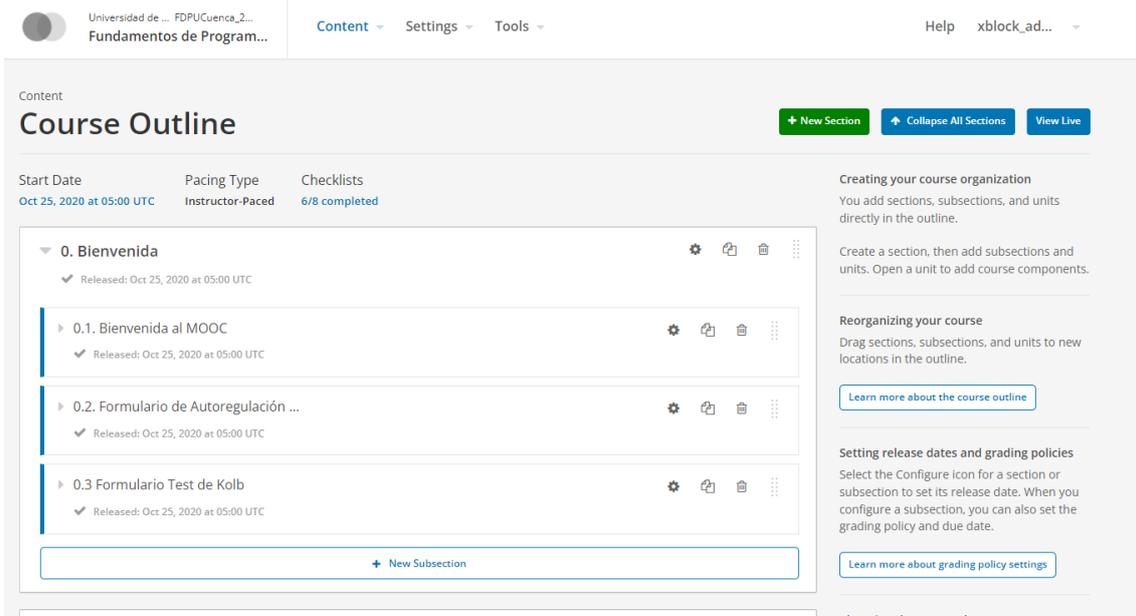


Figura 49. Selección del curso donde se va habilitar el XBlock.

Una vez en el curso, el administrador deberá ir a la pestaña “Tools” en la parte superior de la interfaz (ver **Figura 50**). Luego hay que seleccionar la opción “Advanced Settings”.

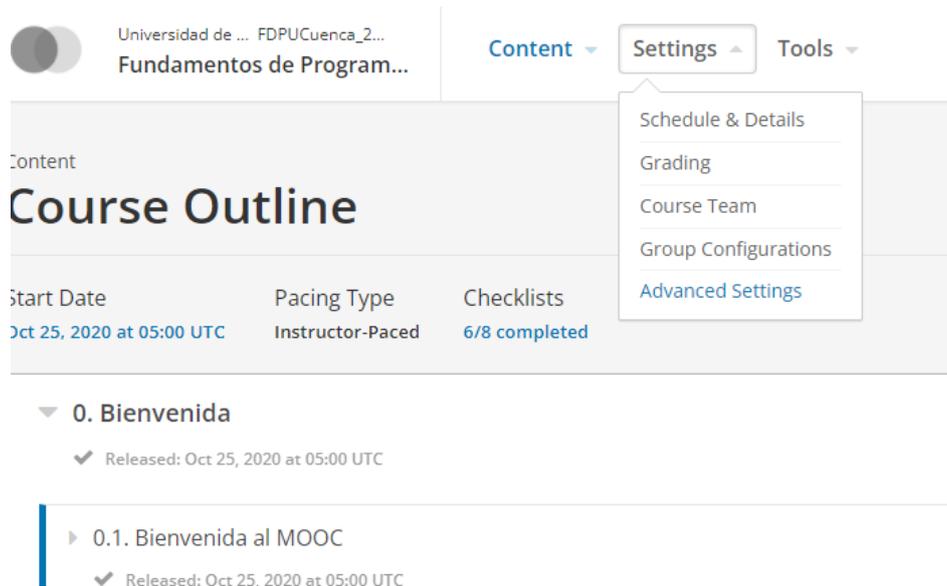


Figura 50. Seleccionar configuración avanzada del curso.

En la interfaz mostrada en la **Figura 51**, el administrador deberá dirigirse a la sección “Advanced Module List” y luego ingresar entre comillas dobles el nombre del componente XBlock que desea habilitar dentro del curso, en este caso el XBlock se llama “xlea”.

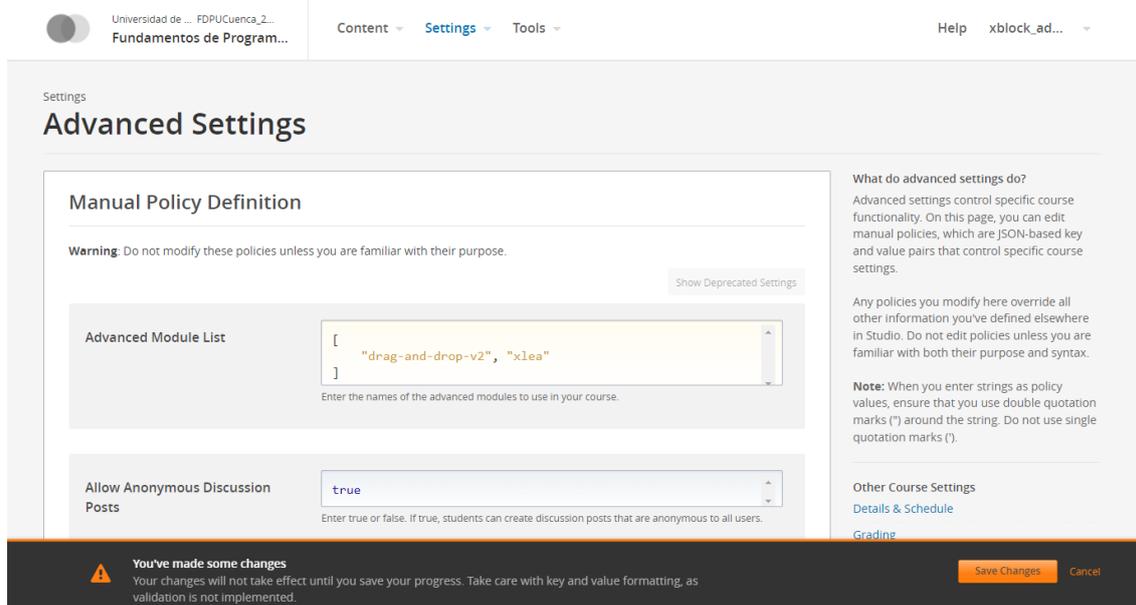


Figura 51. Habilitación del XBlock desarrollado en este trabajo de titulación.

Luego de guardar los cambios, el administrador deberá dirigirse al panel general del curso y crear un nuevo módulo donde el XBlock será mostrado a los estudiantes y maestros. Dentro de este módulo se deberá crear una sección y posteriormente una unidad (ver **Figura 52**), esto se debe a que los XBlocks en Open edX solo pueden instalarse en la unidad más pequeña de un curso.

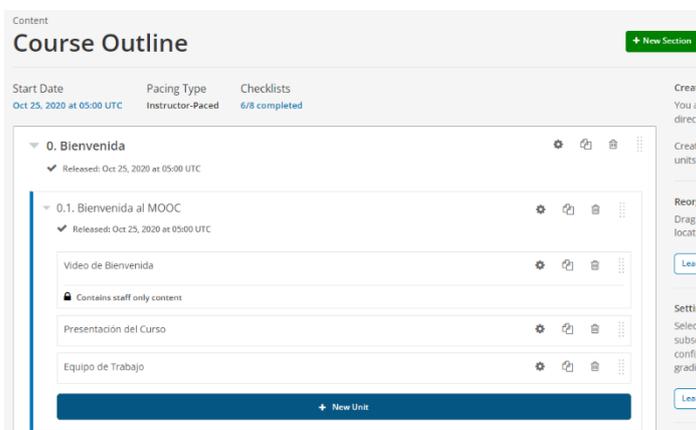


Figura 52. Creación de la unidad donde será mostrado el XBlock dentro del curso.

Dentro de la unidad que contendrá el XBlock se mostrarán 5 opciones de color verde (ver **Figura 53**), el usuario deberá seleccionar la opción “Advanced”.

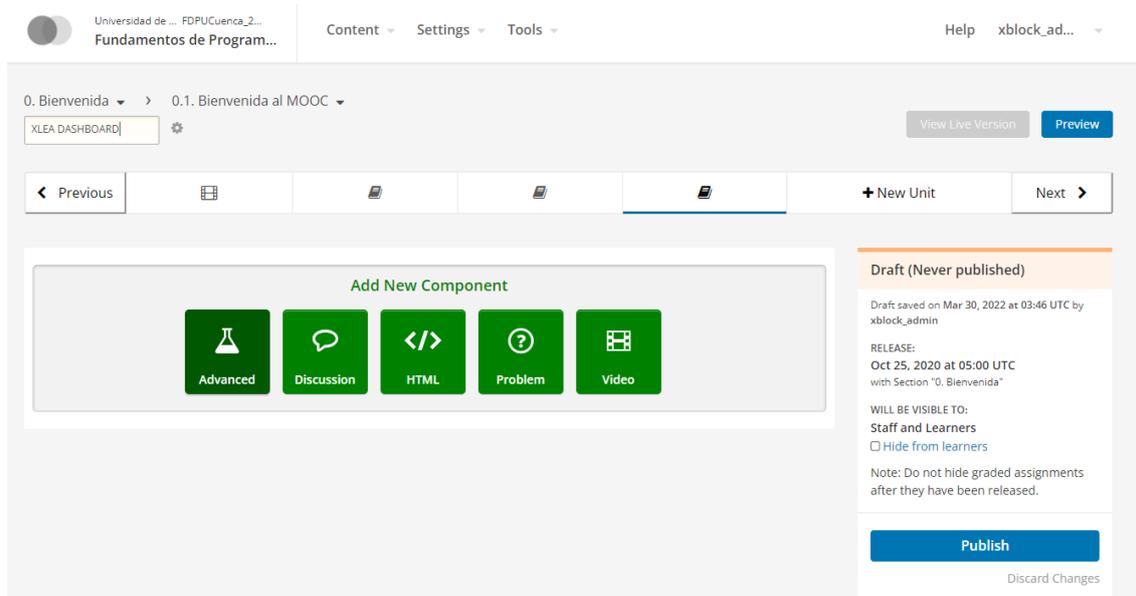


Figura 53. Agregación de nuevo componente avanzado dentro del curso.

Luego, se mostrará una lista de los XBlocks disponibles para mostrarse dentro de la unidad. Aquí el administrador deberá seleccionar el nombre del XBlock desarrollado en este trabajo de titulación (ver **Figura 54**).

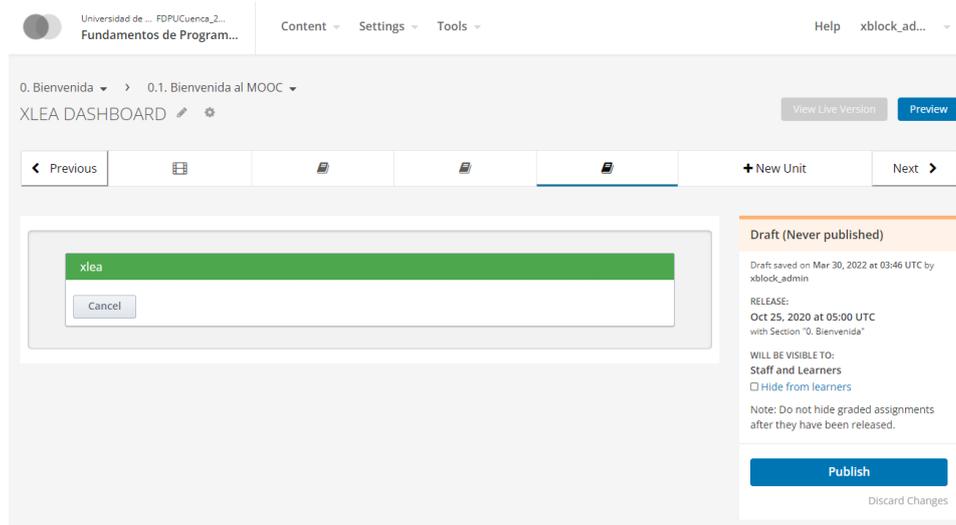


Figura 54. Selección del XBlock a mostrar dentro de la unidad del curso.

Una vez realizado este paso, el administrador deberá guardar los cambios y publicar la unidad con el XBlock para que sea mostrado a estudiantes y maestros dentro del curso, tal como se muestra en la **Figura 55**.

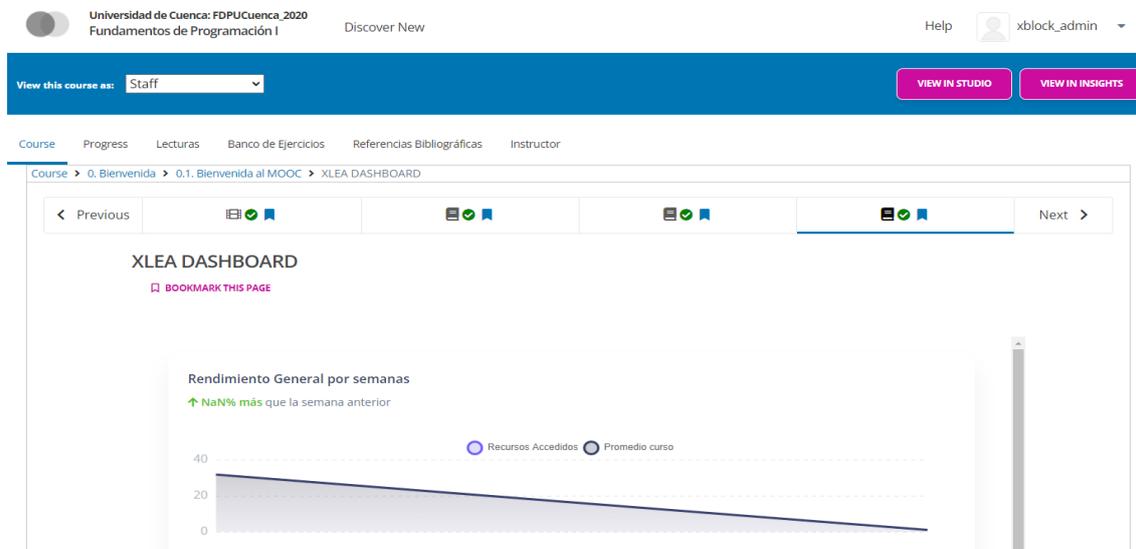


Figura 55. Previsualización del dashboard de la vista del estudiante dentro del curso.

6.5. Evaluación del componente de visualizaciones

Finalmente, en esta sección se introduce al lector el contexto en el que se realizará la evaluación del componente de LA desarrollado en este trabajo de titulación. Para la evaluación del dashboard se hará uso de dos cuestionarios, el primero es EFLA (Scheffel et al., 2017) y el segundo es UEQ (Laugwitz et al., 2008), con el propósito de evaluar las analíticas de aprendizaje y la experiencia de usuario respectivamente.

Se eligió como herramienta de evaluación el cuestionario EFLA, debido a que fue desarrollada y validada para estandarizar la manera en la que se evalúan herramientas de LA. EFLA emplea 4 dimensiones de evaluación de herramientas de LA (ver **Figura 56**), las cuáles se basan en la metodología Group Concept Mapping (GCM, por sus siglas en inglés) (Kane & Trochim, 2009) y están validadas por el Technology Acceptance Model (TAM, por sus siglas en inglés), dichas dimensiones se mencionan a continuación :

Tabla 12. Dimensiones de evaluación de herramientas de LA.

Nota: Tomado de (Scheffel et al., 2017)

Dimensión	Descripción	Indicadores
Datos	Esta variable está relacionada al acceso y la presentación de los datos	Transparencia
		Manipulación
Concientización y Reflexión	Esta variable está relacionada a la proyección a futuro del estado de aprendizaje de los estudiantes.	Comparabilidad
		Recomendación
		Clasificación de actividades
		Comportamiento
Impacto	Esta variable está relacionada a la motivación que sienten los usuarios al usar la herramienta.	Efectividad
		Eficiencia

Figura 56. Dimensiones de evaluación de herramientas de LA.

Nota: Tomado de (Scheffel et al., 2017)

UEQ es un cuestionario que es ampliamente usado para medir la impresión subjetiva que tienen los usuarios sobre la calidad y usabilidad en los productos de software. (Schrepp et al., 2017). En el cuestionario empleado se usa una escala de Likert para medir las reacciones, actitudes y comportamientos de las personas donde 1 es totalmente en desacuerdo y 5 totalmente de acuerdo; se ha seleccionado esta escala debido a que a diferencia de las preguntas de si/no está permite tener una mayor libertad en calificar las respuestas (Nemoto & Beglar, 2014). Los factores empleados para medir la experiencia que tiene el usuario con el producto de software se muestran a continuación:

- **Atractivo:** Impresión general del producto. ¿A los usuarios les gusta o disgusta el producto?
- **Claridad:** ¿Es fácil familiarizarse con el producto? ¿Es fácil aprender a usar el producto?
- **Eficiencia:** ¿Pueden los usuarios resolver sus tareas sin esfuerzo innecesario?
- **Confianza:** ¿Se siente el usuario en control de la interacción?

- **Estímulo:** ¿Es emocionante y motivador usar el producto?
- **Novedad:** ¿Es el producto innovador y creativo? ¿El producto capta el interés de los usuarios?

6.5.1. Selección de los involucrados en la evaluación

La evaluación de las visualizaciones incorporadas en el XBlock será realizada por estudiantes y profesores de la Universidad de Cuenca (muestreo a conveniencia). Los participantes tendrán acceso completo a todas las visualizaciones, tomando en cuenta su rol dentro de la plataforma se proporcionará el cuestionario respectivo. La cantidad de participantes seleccionados son 14, de los cuales 10 son estudiantes de la carrera de Ingeniería en ciencias de la computación que estaban cursando el último año de la carrera y que previamente han tenido una capacitación de que es un MOOC, debido a que han tomado previamente un curso denominado “Como realizar un MOOC”; los 4 profesores que nos participaron en la evaluación de igual manera han tenido experiencia en impartir cursos MOOC.

6.5.2. Aplicación del cuestionario a los evaluadores

En esta subsección se explica la manera en la que se ejecutarán los dos cuestionarios dispuestos por la herramienta de evaluación EFLA. Ambos cuestionarios están dirigidos para estudiantes y docentes tomando en cuenta las necesidades y objetivos de ambos (Scheffel et al., 2017). Las preguntas que son empleadas para medir las variables en función de la percepción obtenida para los estudiantes como los profesores se muestran la **Figura 57**.

**The Evaluation Framework for Learning Analytics
EFLA**

for
LEARNERS

for
TEACHERS

DATA											
For this LA tool it is clear what data is being collected											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
For this LA tool it is clear why the data is being collected											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○

DATA											
For this LA tool it is clear what data is being collected											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
For this LA tool it is clear why the data is being collected											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○

AWARENESS & REFLECTION											
This LA tool makes me aware of my current learning situation											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool makes me forecast my possible future learning situation given my (un)changed behaviour											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool stimulates me to reflect on my past learning behaviour											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool stimulates me to adapt my learning behaviour if necessary											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○

AWARENESS & REFLECTION											
This LA tool makes me aware of my students' current learning situation											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool makes me forecast my students' possible future learning situation given their (un)changed behaviour											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool stimulates me to reflect on my past teaching behaviour											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool stimulates me to adapt my teaching behaviour if necessary											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○

IMPACT											
This LA tool stimulates me to study more efficiently											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool stimulates me to study more effectively											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○

IMPACT											
This LA tool stimulates me to teach more efficiently											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○
This LA tool stimulates me to teach more effectively											
strongly disagree	1	2	3	4	5	6	7	8	9	10	strongly agree
○	○	○	○	○	○	○	○	○	○	○	○

EFLA scoring per stakeholder group										
Step 1: Calculate the average value for each item based on the answers given for that item.										
Step 2: Calculate the average value for each dimension based on the average of its items.										
Step 3: Calculate the dimensional scores by rounding the results of $((x-1)/9)*100$ where x is the average value of a dimension.										
Step 4: Calculate the overall EFLA score by taking the average of the three dimensional scores.										

Figura 57. Cuestionario para la evaluación de herramientas de LA, estudiante y docente.

Nota: Tomado de (Scheffel et al., 2017)

Estas preguntas deben responderse en una escala del 1 al 10, siendo 10 el valor de mayor importancia mientras que el valor de 1 es el valor de menor importancia.

Además, para obtener cada uno de los valores de cada dimensión se define la ecuación (1).

$$Dimension\ Score = \left(\frac{x-1}{9}\right) * 100 \quad (1)$$

Para el cuestionario UEQ se han realizado las preguntas que provee el cuestionario en base a sus factores, obteniendo las preguntas que será usadas para la evaluación de la herramienta XLEA, las mismas que se encuentran a continuación:

Tabla 13. Estructura del cuestionario UEQ.

Factor	Definición	Pregunta
Atractivo	El producto se ve atractivo, ameno, amigable y placentero	Me pareció atractiva la herramienta
Eficiencia	Puedo realizar mis tareas con el producto de manera rápida, eficiente y pragmática. La interfaz de usuario se ve organizada	Las figuras se encuentran organizadas de una manera fácil de entender
Claridad	El producto es fácil de entender, claro, simple y fácil de aprender	Tuve que aprender otras cosas para usar la herramienta
Confianza	La interacción con el producto es predecible, segura y cumple con mis expectativas. El producto me ayuda a realizar mis tareas	El producto me ayuda a tener un control sobre mi aprendizaje
Estímulo	El uso del producto es interesante, excitante y motivador	La herramienta captó mi interés
Novedad	El producto es innovador, inventivo y tiene un diseño creativo	La herramienta motiva a utilizarla de nuevo

6.6. Resultados de la evaluación a estudiantes empleando EFLA

UCUENCA

El uso del cuestionario EFLA para los estudiantes tiene como propósito evaluar las analíticas de aprendizaje implementadas en el dashboard que es mostrado a todos los estudiantes por medio del Xblock.

La **Figura 58 - 59** representan las respuestas otorgadas por los estudiantes respecto a la dimensión relacionada a los datos, específicamente a su acceso y representación. En las respuestas otorgadas queda claro que la mayoría de ellos están entendiendo que datos se están recolectando y el por qué lo están recolectando de manera que queda claro los indicadores de transparencia y la manipulación.

Esta herramienta de LA, queda claro qué datos se recolectan

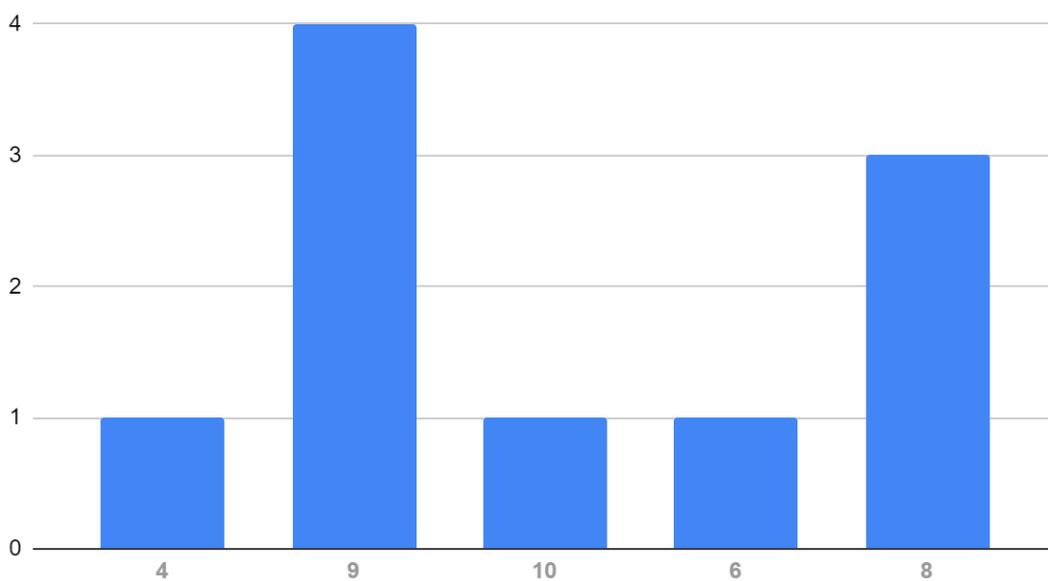


Figura 58. Respuestas a la primera pregunta del EFLA para estudiantes.

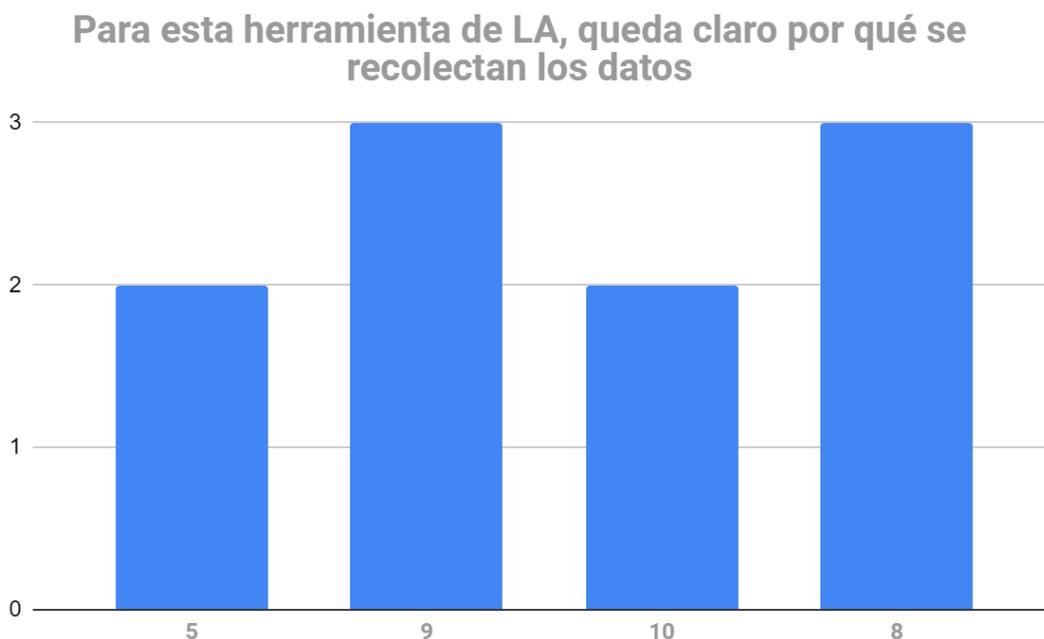


Figura 59. Respuestas a la segunda pregunta del EFLA para estudiantes.

Desde la **Figuras 60 a 63** representan los resultados obtenidos de los estudiantes con respecto a la dimensión de Concientización y Reflexión, la cual está relacionada a la proyección a futuro del aprendizaje que están teniendo los estudiantes. Los resultados reflejan que la herramienta si muestra de manera satisfactoria la situación actual de su aprendizaje; de igual manera casi más de la mitad de los estudiantes respondieron que les estimula a adaptar su comportamiento de enseñanza en caso de ser necesario; y un poco menos de la mitad de los estudiantes comentaron que la herramienta no les permitía pronosticar la posible situación de aprendizaje con su comportamiento actual. Dado esto se puede mencionar que la herramienta sufre un poco en LA que permitan pronosticar la situación futura del aprendizaje de los estudiantes.

Esta herramienta de LA me hace consciente de mi situación de aprendizaje actual

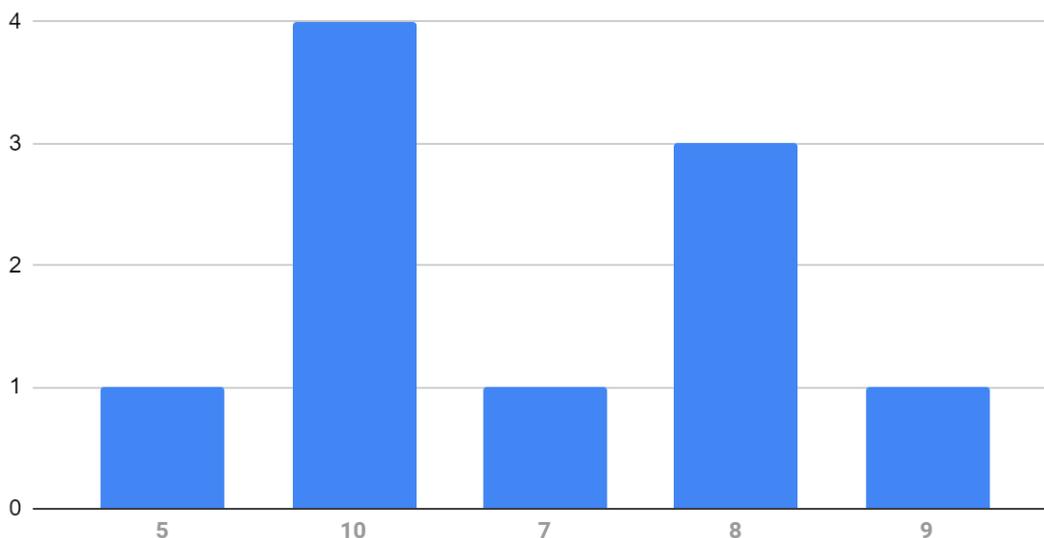


Figura 60. Respuestas a la tercera pregunta del EFLA para estudiantes.

Esta herramienta de LA me hace pronosticar mi posible futuro situación de aprendizaje dada mi comportamiento (no) cambiado

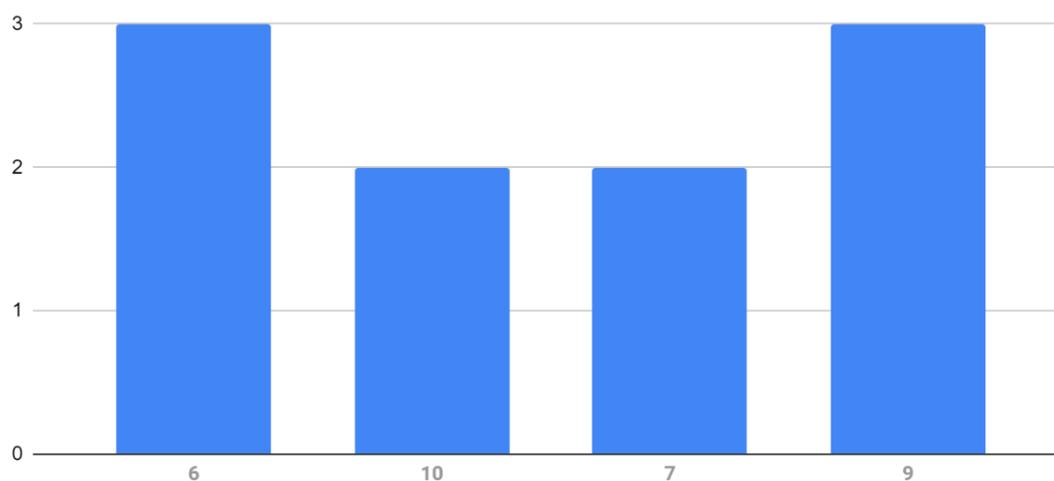


Figura 61. Respuestas a la cuarta pregunta del EFLA para estudiantes.

Esta herramienta de LA me estimula a reflexionar sobre mi comportamiento de enseñanza pasado

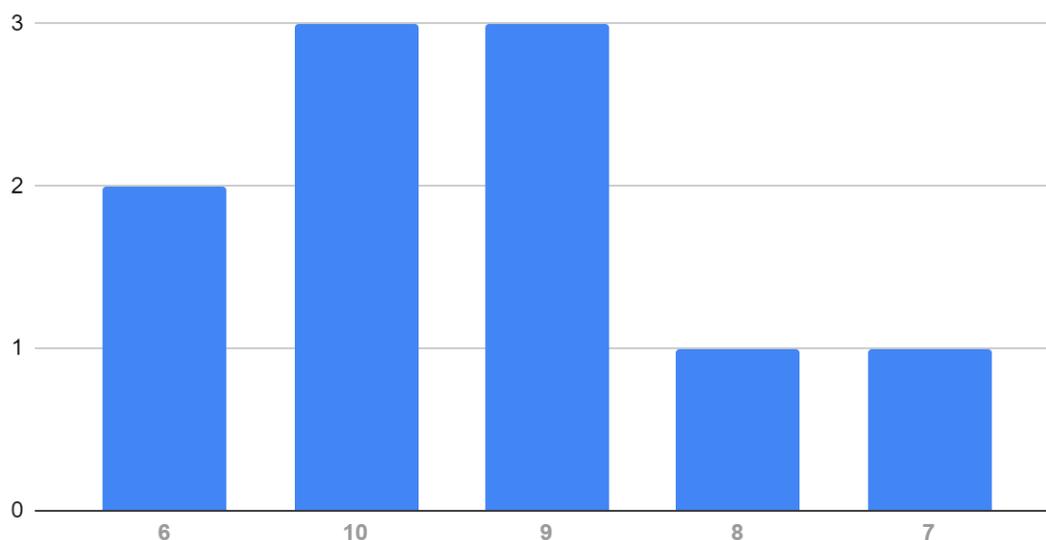


Figura 62. Respuestas a la quinta pregunta del EFLA para estudiantes.

Esta herramienta de LA me estimula a adaptar mi comportamiento en la enseñanza si es necesario

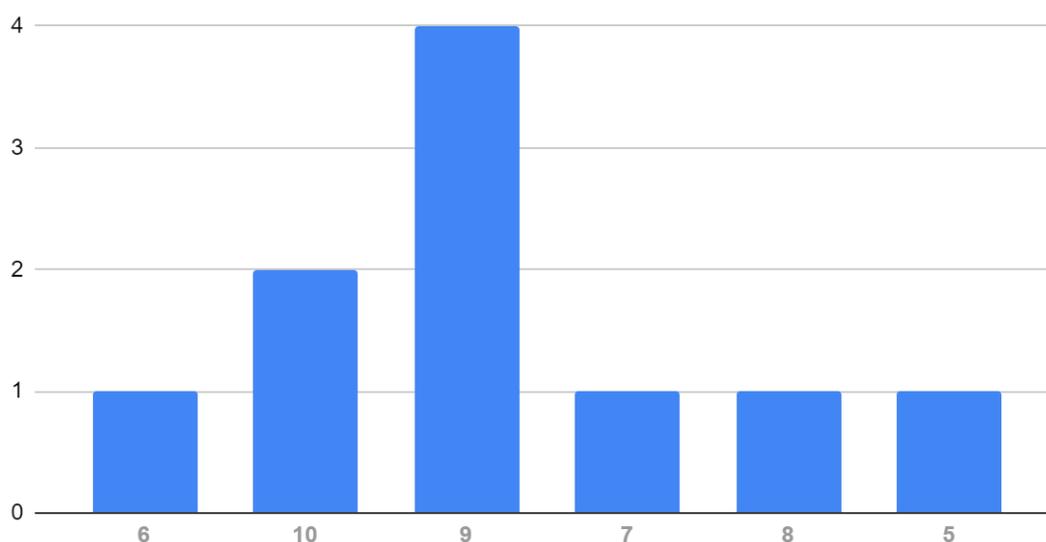


Figura 63. Respuestas a la sexta pregunta del EFLA para estudiantes.

UCUENCA

Las **Figuras 64 - 65** representan los resultados obtenidos de los estudiantes con respecto a la dimensión del Impacto, la cual está relacionada a la motivación que sienten los usuarios al usar la herramienta. Los resultados reflejan que más de la mitad de los estudiantes se sintieron estimulados a estudiar de manera más eficiente y efectiva, por lo que los estudiantes aseguran volver a usar la herramienta y existe una pequeña cantidad de estudiantes que no se sintieron tan estimulados como la mayoría de los estudiantes.

Esta herramienta de LA me estimula a estudiar de manera más eficiente

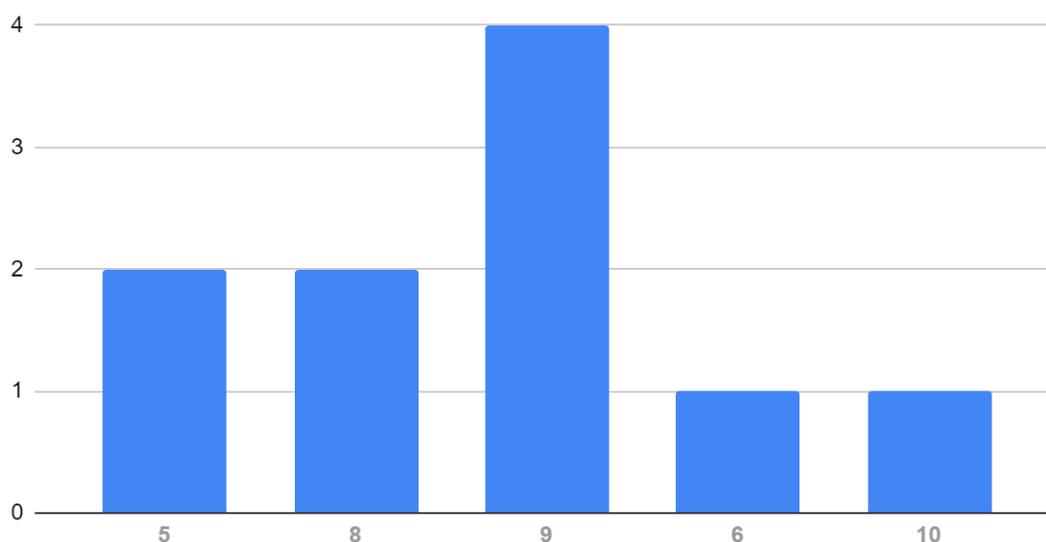


Figura 64. Respuestas a la séptima pregunta del EFLA para estudiantes.

Esta herramienta de LA me estimula a estudiar de manera más efectiva

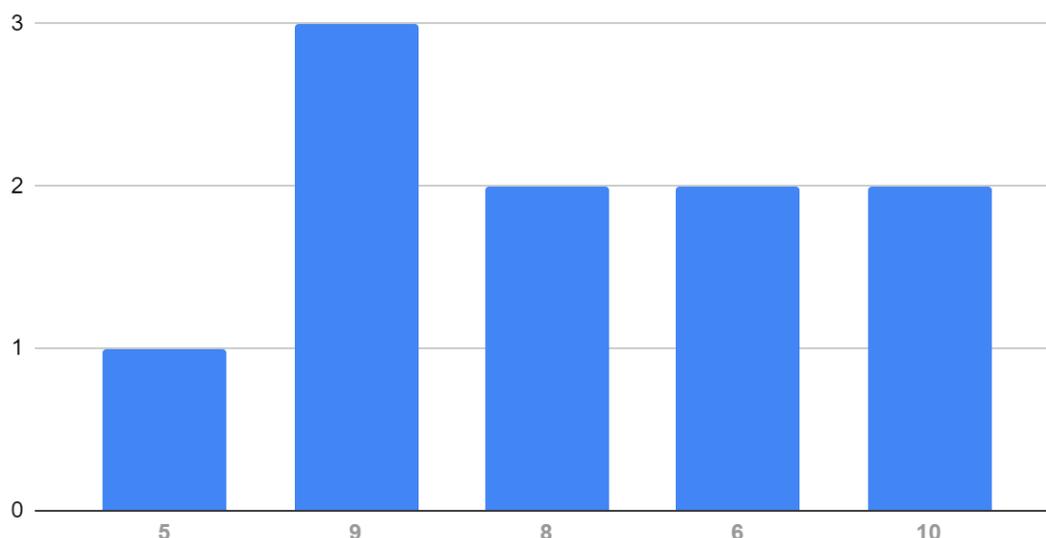


Figura 65. Respuestas a la octava pregunta del EFLA para estudiantes.

Del previo análisis de los indicadores que son mostrados en el dashboard de LA, se ha analizado la suma de los valores de los participante y obtenido el EFLA Score mostrados en la **Tabla 14**. En el análisis del EFLA Score se obtuvieron valores mayores a 60 en en todas las dimensionalidades analizadas, indicando que las LA implementadas para los estudiantes son de ayuda para tener un seguimiento del progreso que están teniendo dentro de un curso MOOC.

Tabla 14. Resultados de aplicar el cuestionario EFLA en los estudiantes.

Dimensión	Suma	EFLA Score
Datos	81.5	79.44
Concientización y Reflexión	83.5	81.67
Impacto	80	77.78

6.7. Resultados de la evaluación a docentes usando EFLA

El uso del cuestionario EFLA para los docentes tiene como propósito evaluar las analíticas de aprendizaje implementadas en el dashboard que es mostrado a todos los docentes por medio del XBlock.

UCUENCA

La **Figura 66 - 67** se muestran la frecuencia de las respuestas otorgadas por los docentes respecto a la dimensión relacionada a los datos, específicamente a su acceso y representación. En las respuestas se puede analizar que los docentes consideran que tienen claro cuáles son los datos que se están recolectando y el por qué se están recolectando quedando claro los indicadores de transparencia y la manipulación.

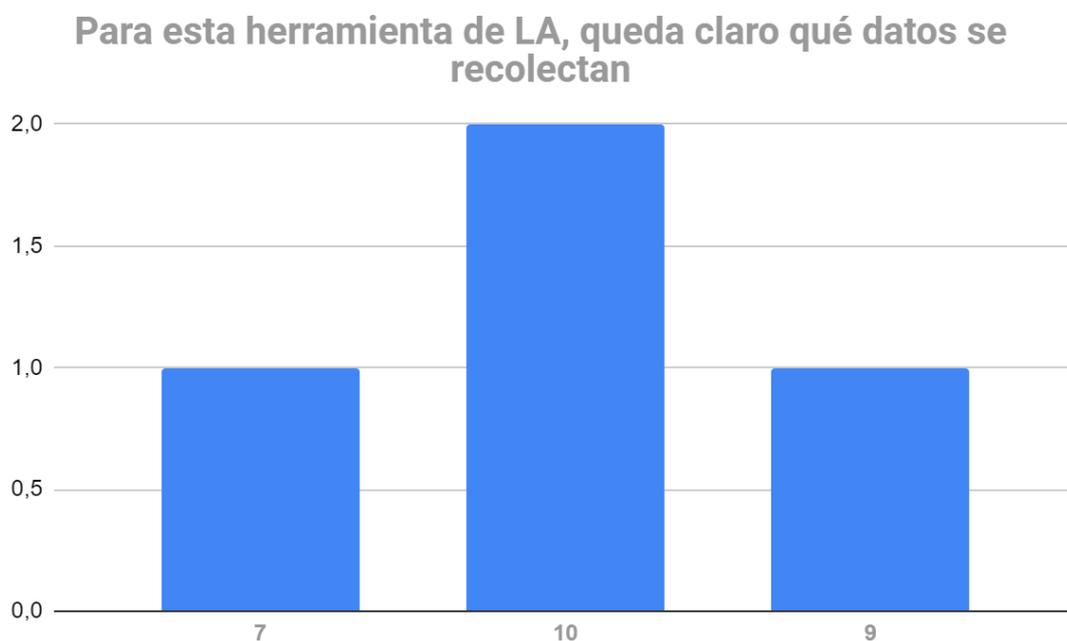


Figura 66. Respuestas a la primera pregunta del EFLA para profesores.

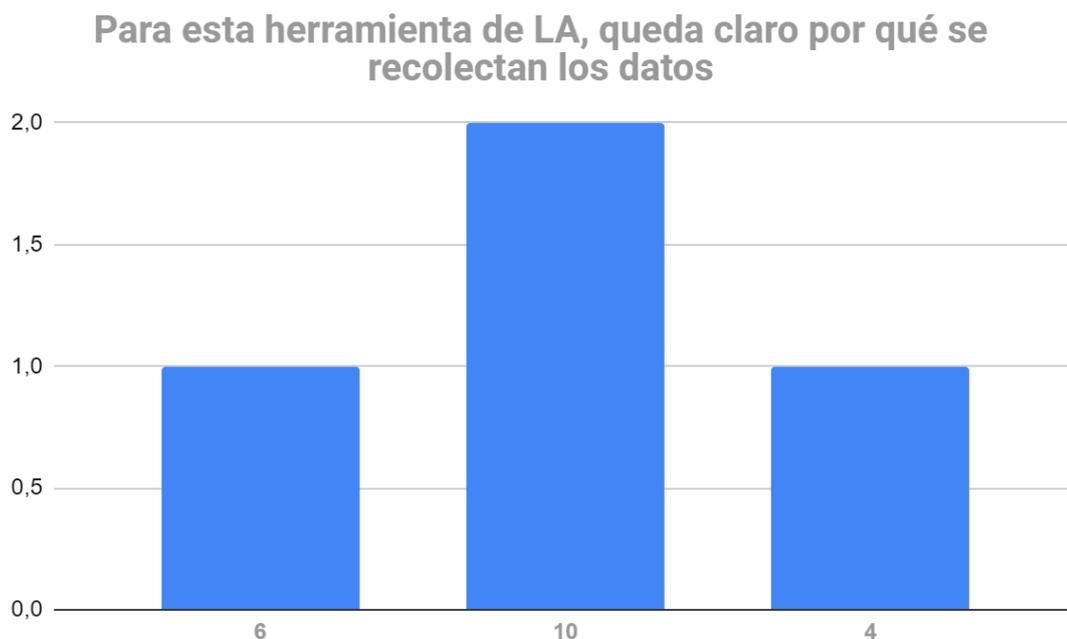


Figura 67. Respuestas a la segunda pregunta del EFLA para profesores.

Desde las **Figuras 68 a 71** representan los resultados obtenidos de los docentes con respecto a la dimensión de Concientización y Reflexión, la cual está relacionada a la proyección a futuro del aprendizaje que están teniendo los estudiantes. Los resultados reflejan que la herramienta muestra de manera satisfactoria el estado actual y pasado que han tenido los alumnos en el aprendizaje; con respecto al posible escenario de los alumnos dado un cambio en el comportamiento los resultados muestran que la mitad de los docentes estaba de acuerdo con esto mientras que la otra mitad no estaba de acuerdo con lo antes mencionado, es decir se tiene una opinión dividida. Dado este resultado en el dashboard se debe de mejorar las LA para mostrar los escenarios que tienen los alumnos cuando tienen un cambio en su comportamiento.

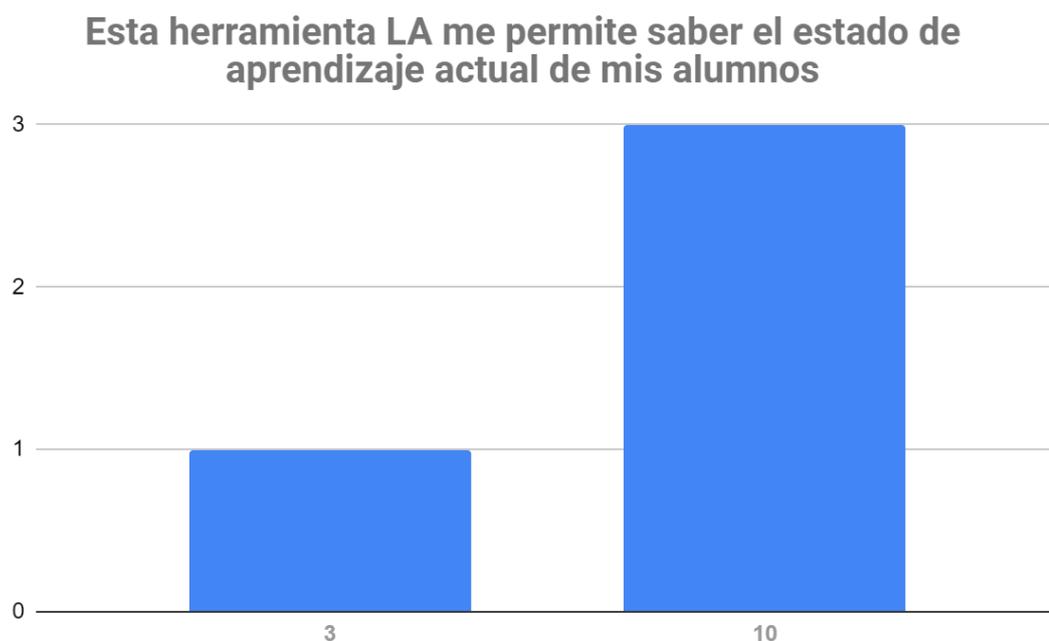


Figura 68. *Repuestas a la tercera pregunta del EFLA para profesores.*

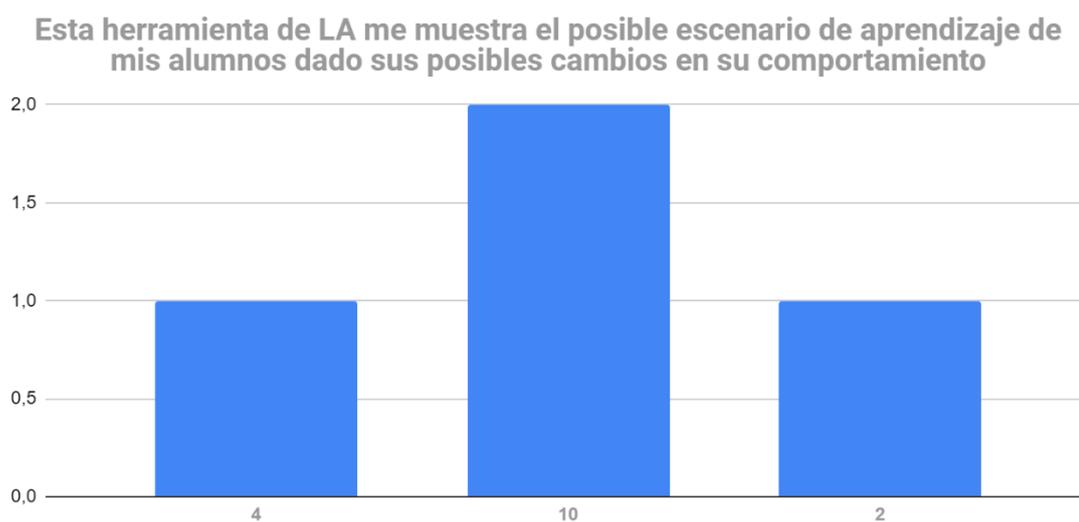


Figura 69. *Repuestas a la cuarta pregunta del EFLA para profesores.*



Figura 70. Respuestas a la quinta pregunta del EFLA para profesores.

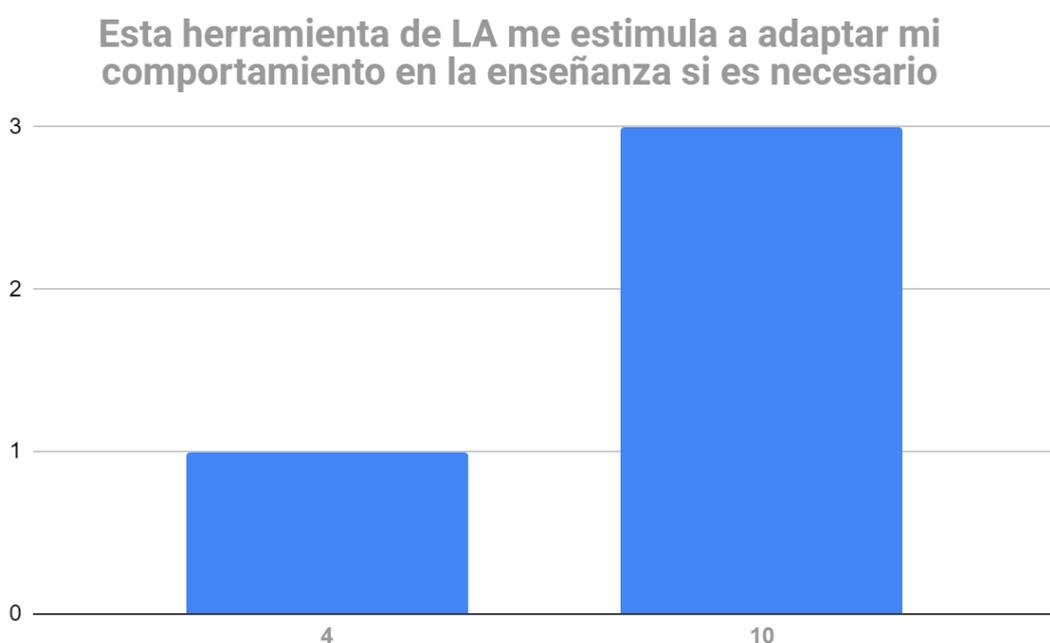


Figura 71. Respuestas a la sexta pregunta del EFLA para profesores.

UCUENCA

Las **Figuras 72 - 73** representan los resultados obtenidos de los docentes con respecto a la dimensión del Impacto, la cual está relacionada a la motivación que sienten al usar la herramienta. Los resultados reflejan que más de la mitad de los docentes se sintieron estimulados a enseñar de manera más eficiente y efectiva, por lo que los docentes volverían usar la herramienta y existe una pequeña cantidad de docentes que no se sintieron tan estimulados como la mayoría.

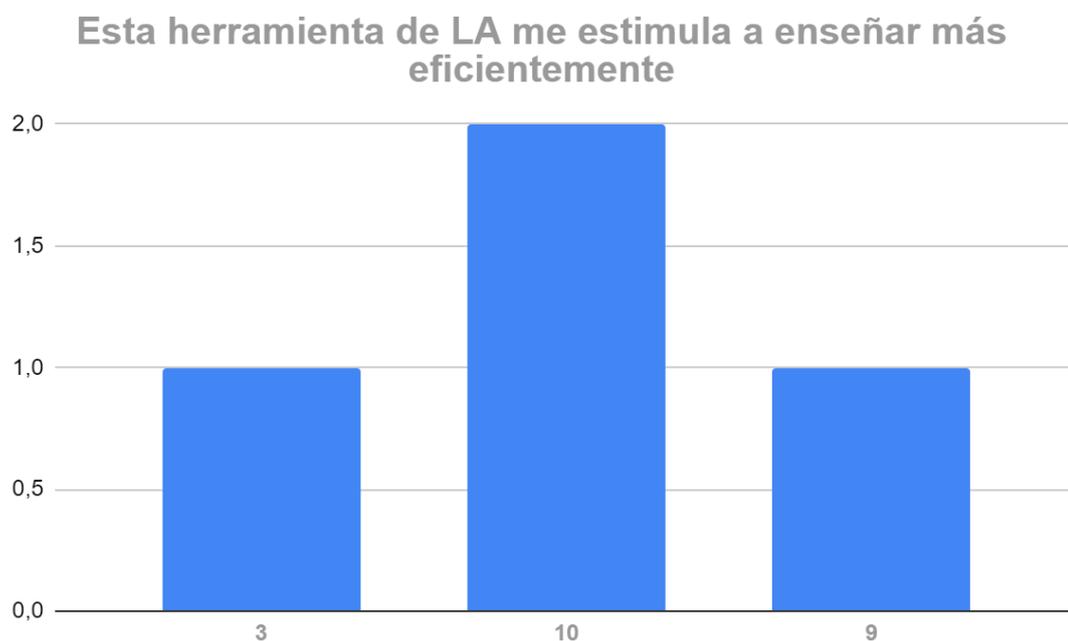


Figura 72. Respuestas a la séptima pregunta del EFLA para profesores.

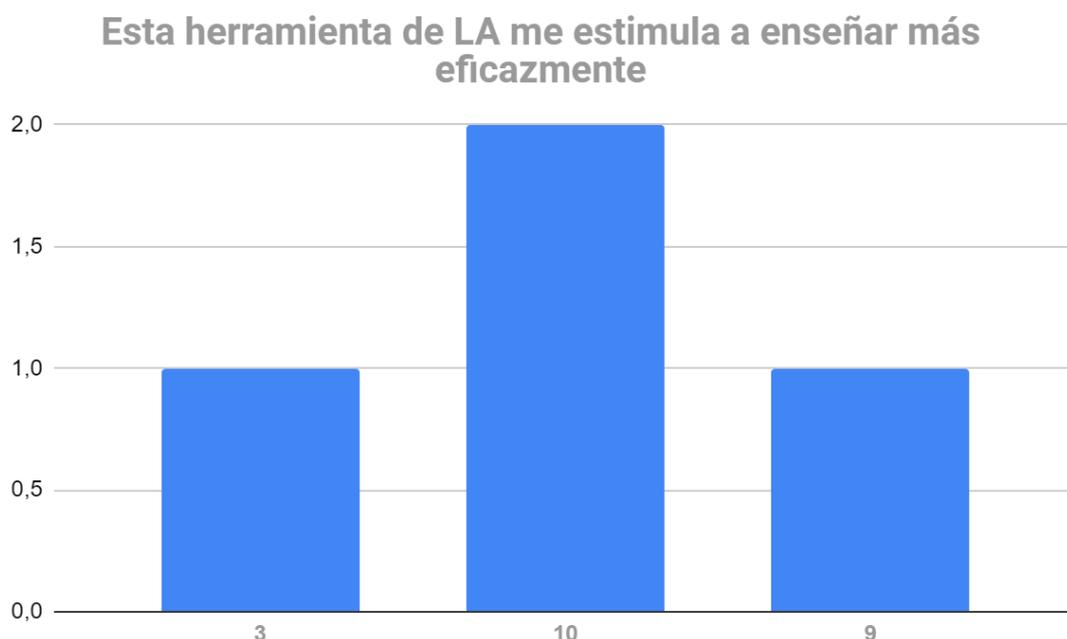


Figura 73. Respuestas a la octava pregunta del EFLA para profesores.

Del análisis anterior de los indicadores del dashboard, se ha realizado un análisis de la suma de sus valores y el valor obtenido del EFLA Score mostrado en el **Tabla 15**. En el análisis del EFLA Score se obtuvieron valores mayores a 60 en cada una de las dimensionalidades analizadas por este cuestionario, lo que indica que las LA implementadas para el dashboard del docentes están siendo de ayuda en el seguimiento del aprendizaje de sus estudiantes.

Tabla 15. Resultados de aplicar el cuestionario EFLA en los docentes.

Dimensión	Suma	EFLA Score
Datos	33	80.55
Concientización y Reflexión	31.75	77.08
Impacto	32	77.78

6.8. Resultados de la evaluación de experiencia de usuario

Se utilizó el cuestionario UEQ en el cual participaron estudiantes y docentes por igual, cuyo objetivo es analizar la calidad en la experiencia de usuario que provee el dashboard desarrollado en este trabajo de titulación.

UCUENCA

En la **Figura 74** se presentan las respuestas relacionadas a qué tan atractivo se percibe el dashboard de analíticas de aprendizaje. Más de la mitad de los participantes están de acuerdo a que la herramienta genera un grado de atención en sus usuarios, reflejando el interés que sienten por usar el dashboard en una primera instancia.

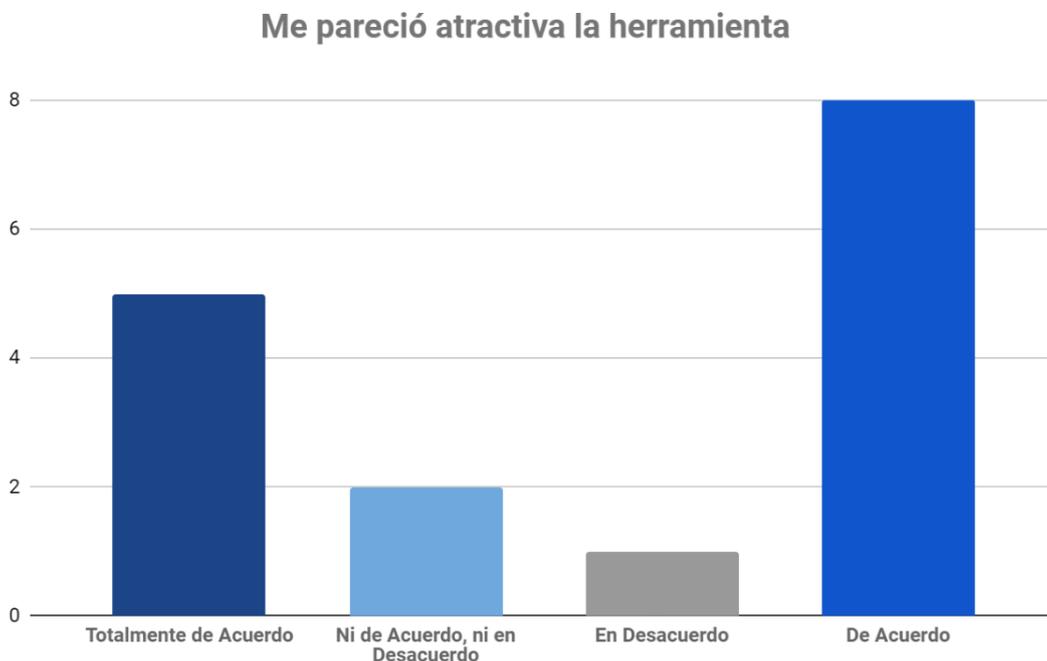


Figura 74. Respuestas a la primera pregunta del UEQ.

Cerca del 80% de los participantes se encuentran en acuerdo en que el dashboard presenta las visualizaciones de forma organizada y fácil de entender, como se puede apreciar en la **Figura 75**.

Las figuras se encuentran organizadas de una manera facil de entender

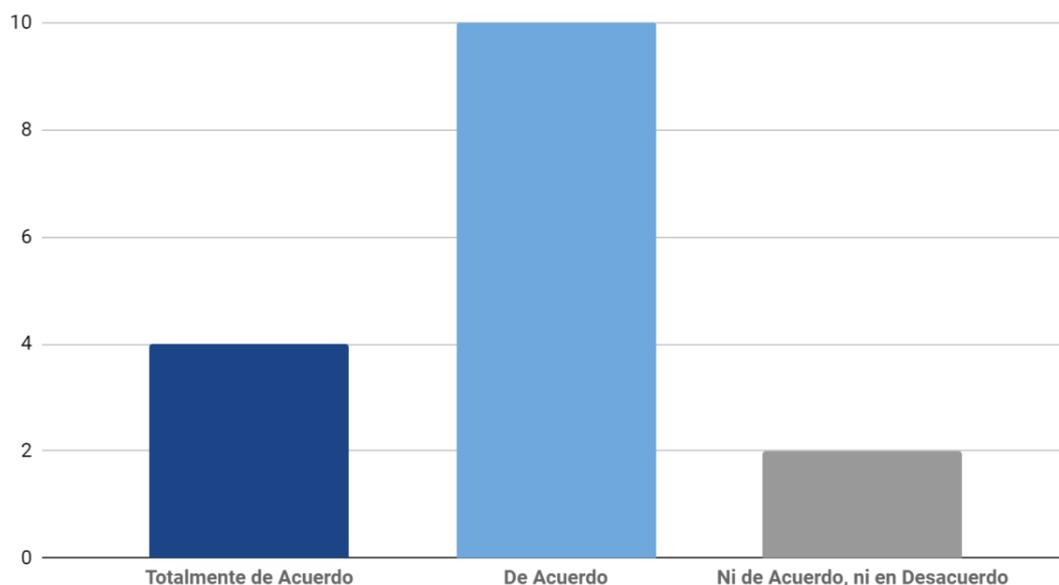


Figura 75. Respuestas a la segunda pregunta del UEQ.

En la tercera pregunta del cuestionario (ver **Figura 76**), debido a que la mayoría de los participantes han tenido experiencia en el uso de dashboards y de la plataforma Open edX, expresan que no hubo necesidad de aprender alguna otra herramienta.

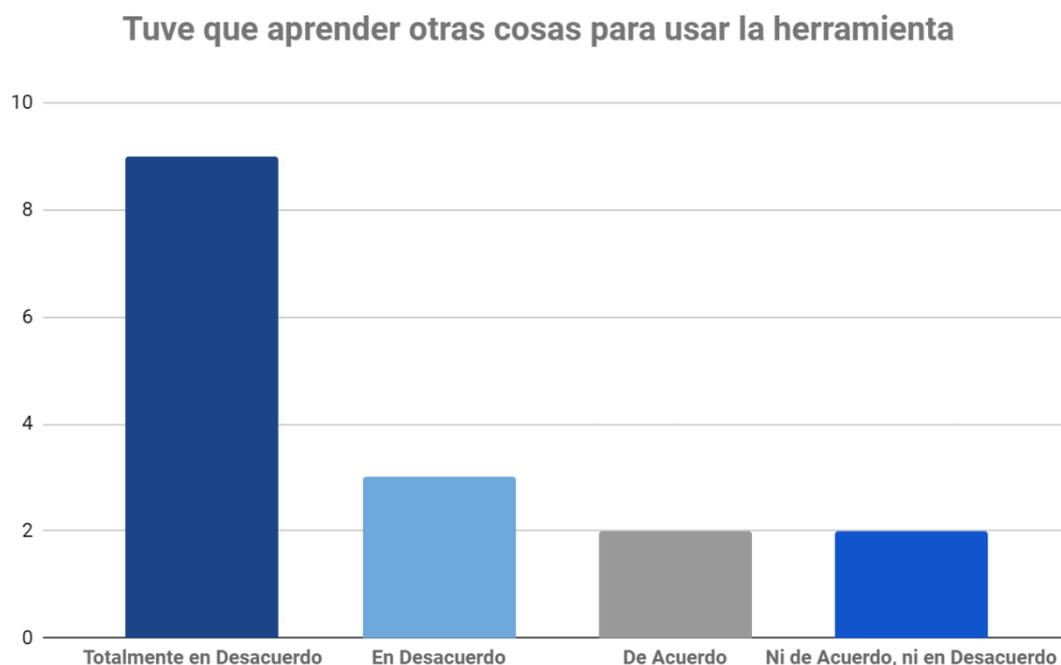


Figura 76. Respuestas a la tercera pregunta del UEQ.

En la **Figura 77** se puede observar que el 70% de los participantes están de acuerdo con que el dashboard XLEA les permite tener control sobre su proceso de aprendizaje.

Esta herramienta me ayuda a mi a tener un control sobre mi aprendizaje

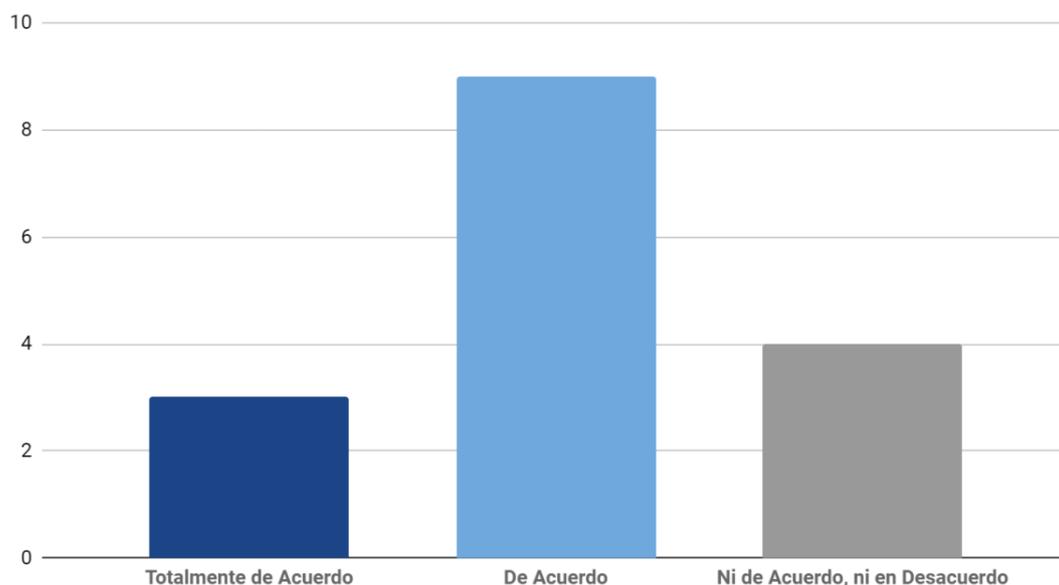


Figura 77. Respuestas a la cuarta pregunta del UEQ.

En la quinta pregunta, aproximadamente el 70% de los participantes indican haber mantenido su interés luego de usar el componente de analíticas de aprendizaje. Esto sugiere, que existen participantes que luego de haber usado el dashboard perdieron interés, como se observa en la **Figura 78**.

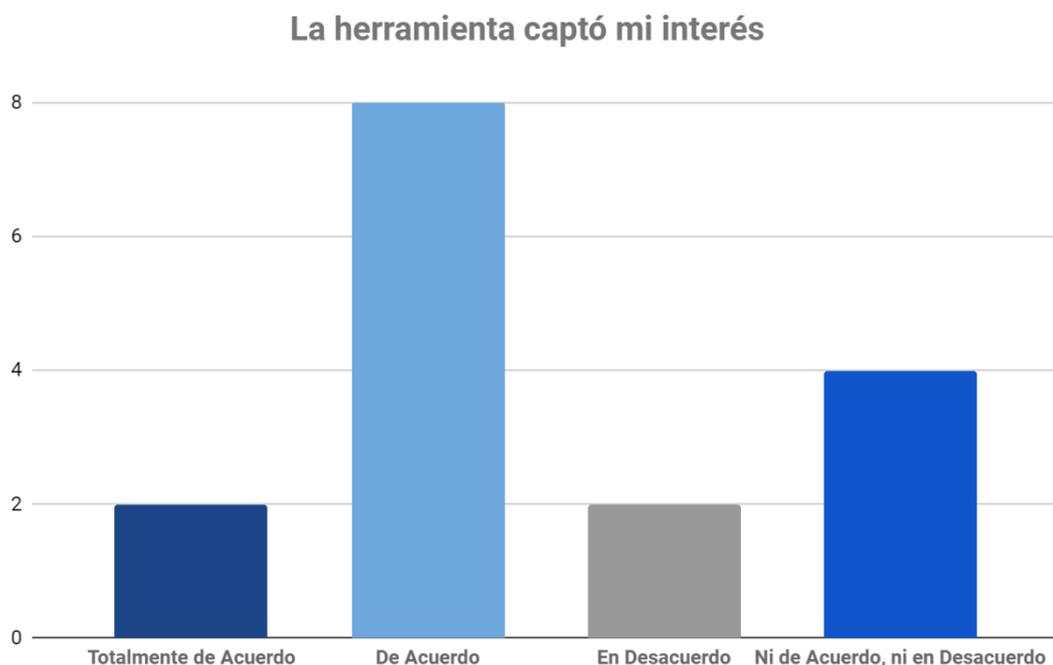


Figura 78. Respuestas a la quinta pregunta del UEQ.

Finalmente, en la sexta pregunta se observa que el 60% de los participantes están dispuestos a seguir utilizando la herramienta, mientras que el 40% restante aún tienen dudas (ver **Figura 79**). Esto se puede deber al margen de mejora que existe en las visualizaciones del dashboard según la retroalimentación de los participantes, tema que se será tratado en la sección de conclusiones y trabajos futuros de este capítulo.

La herramienta motiva a utilizarla de nuevo

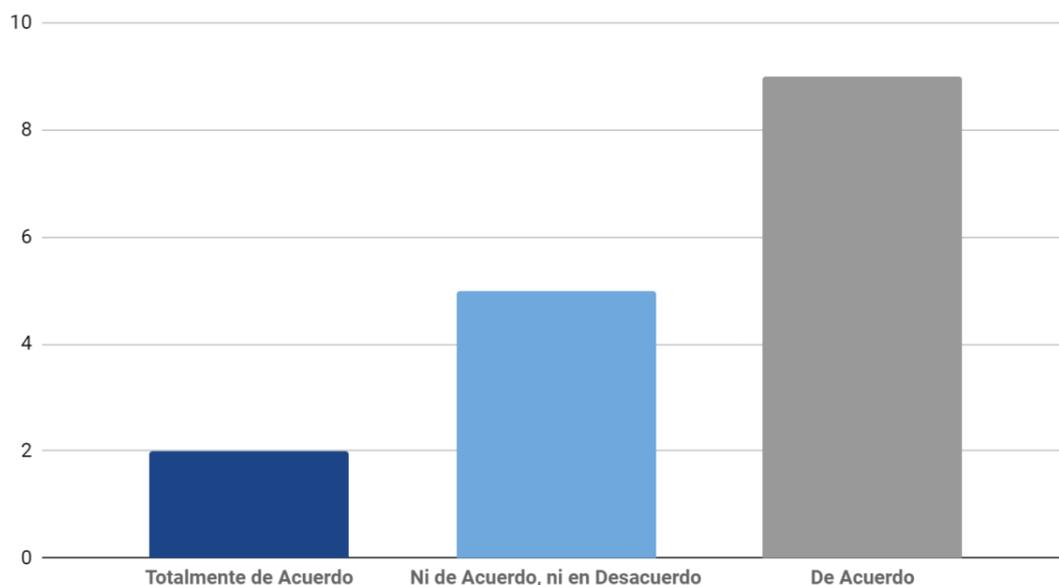


Figura 79. Respuestas a la sexta pregunta del UEQ.

6.9. Conclusiones

En este capítulo se ejecutan las etapas de implementación y evaluación del XBlock de la metodología usada en este trabajo de titulación. Se implementaron las visualizaciones diseñadas en el capítulo 5, se revisó la arquitectura de los componentes de Open edX, así también, se presenta la arquitectura del componente de LA desarrollado. Las funcionalidades que brinda XLEA permiten a estudiantes y maestros tener una vista general de las interacciones con los recursos del curso, medir y comparar logros conseguidos de un estudiante con la media del curso, además de analizar el tiempo dedicado por cada estudiante en periodos semanales.

En la vista del estudiante se muestran gráficas de línea que representa el rendimiento general del estudiante por semanas. Los indicadores utilizados están relacionados a lecturas, videos y problemas. De la misma manera, para la vista del docente se diseñaron visualizaciones en base al promedio de las interacciones de todos los estudiantes del curso, cuyos indicadores se relacionan con la cantidad de recursos completados por semanas de todos los estudiantes, separándolos en lecturas, videos, foros y problemas.

El XBlock se desplegó a producción en colaboración con CEDIA en la dirección pública <https://test-xlearning.cedia.edu.ec>. Los datos de los cursos procesados por el algoritmo de LA fueron dispuestos por la Universidad de Cuenca. Luego, se realizó la evaluación del XBlock con el cuestionario EFLA, la cual es una herramienta desarrollada y validada para estandarizar la manera en la que se evalúan nuevas herramientas para dar soportes a las LA. EFLA mide y compara la calidad del componente XBlock empleando 4 dimensiones de evaluación para herramientas de LA, las cuáles se basan en la metodología GCM. Las dimensiones están relacionadas con los datos, la concientización, reflexión y el impacto. Adicionalmente, se realizó otra evaluación acerca de la experiencia de usuario aplicando el cuestionario UEQ al XBlock desarrollado en este trabajo de titulación, esto con el propósito de medir la calidad de las visualizaciones incorporado en el componente de LA.

Para la evaluación del dashboard XLEA se escogieron estudiantes y maestros de un curso de postgrado de la Universidad de Cuenca. Los participantes tuvieron acceso completo a todas las visualizaciones tomando en cuenta su rol dentro de la plataforma.

Finalmente, luego de evaluar el dashboard, los resultados evidencian que aproximadamente el 80% de los participantes perciben al dashboard XLEA como un estímulo positivo para redirigir su comportamiento dentro del curso. Enfocándonos en los resultados de EFLA aplicado a estudiantes, las 4 dimensiones de evaluación mantienen un puntaje por arriba de los 77/100 puntos, esto demuestra que la herramienta XLEA presenta de manera sencilla y clara los datos en sus visualizaciones, de igual manera, los estudiantes se sienten motivados a reflexionar acerca de su comportamiento y proceso de aprendizaje. De igual manera los resultados del EFLA aplicado a docentes, se observa que las 4 dimensiones de evaluación mantienen valores por encima de los 77/100 puntos, destacando con 80 puntos la dimensión de datos. Este resultado nos indica que los docentes entienden de qué manera se recopilan y muestran los datos acerca de las interacciones de los estudiantes, permitiéndoles reflexionar acerca de su método de enseñanza y en cómo pueden mejorar en el proceso. Adicionalmente, algunos participantes expresaron sus puntos de vista relacionados a futuras mejoras de XLEA, las cuales serán abordadas en la sección de Recomendaciones del Capítulo 7. Estas recomendaciones se enfocan en la manera en la que se describen las funcionalidades de la herramienta, debido a que no todos los participantes tenían experiencia previa al uso de dashboards, lo cual creemos que es válido y se deberá mejorar en futuras versiones.

Capítulo 7: Conclusiones, Recomendaciones y Trabajos futuros

7.1. Introducción

Finalmente, en este capítulo se presenta al lector las conclusiones obtenidas en este trabajo de titulación. Este capítulo se divide en 3 secciones comenzando con las conclusiones a partir de los objetivos planteados, recomendaciones obtenidas y en la última parte se presenta los trabajos futuros que pueden surgir a partir de esta.

7.2. Conclusiones

El motivo principal de este trabajo de tesis fue el desarrollo de un componente de LA que permita a docentes y estudiantes conocer acerca de su proceso de aprendizaje en un MOOC de Open edX. Para conseguirlo se debieron cumplir los siguientes objetivos:

- (1) Desarrollar un análisis exploratorio del comportamiento de los estudiantes para determinar variables y secuencias de aprendizaje realizadas por los estudiantes.
- (2) Diseñar un Dashboard de visualizaciones para profesores y estudiantes que den cuenta del comportamiento de los estudiantes.
- (3) Implementar un XBlock para Open edX para visualizar el comportamiento de los estudiantes en estos cursos.
- (4) Evaluar la usabilidad de las visualizaciones desarrolladas en el XBlock.

En el capítulo 4 se logra cumplir el primer objetivo de este trabajo de titulación mediante técnicas como EDA y métodos de ML con el propósito de (1) determinar variables y secuencias de aprendizaje realizadas por los estudiantes. Los indicadores seleccionados se centran en la frecuencia de las actividades realizadas por los estudiantes, las secuencias y patrones de aprendizaje, y el rendimiento del estudiante a través del tiempo. Del análisis del comportamiento de los estudiantes se obtuvieron tres grupos de estudiantes, los cuales se identificaron como estudiantes estratégicos, estudiantes de muestro y estudiantes comprensivos.

Con el objetivo de (2) diseñar un dashboard de visualizaciones para estudiantes y docentes, en el capítulo 5 se seleccionan aquellas visualizaciones que se enfoquen en los objetivos y necesidades de estudiantes y maestros, en las fuentes de datos, retroalimentación y autorregulación del aprendizaje y en la facilidad para medir el proceso de aprendizaje. La mayoría de las visualizaciones ofrecen a estudiantes y maestros una perspectiva acerca del estado general del curso a través del promedio de los datos analizados por semanas.

UCUENCA

En el capítulo 6, se realiza una revisión de la arquitectura de Open edX con el objetivo de (3) implementar un XBlock de LA para visualizar el comportamiento de los estudiantes dentro de un curso. En la vista del estudiante se muestran gráficas de línea que representa el rendimiento general del estudiante por semanas. De la misma manera, para la vista del docente se diseñaron visualizaciones en base al promedio de las interacciones de todos los estudiantes del curso. Finalmente, se realizó la (4) evaluación del XBlock con el cuestionario EFLA, la cual es una herramienta desarrollada y validada para estandarizar la manera en la que se evalúan nuevas herramientas para dar soportes a las LA. EFLA mide y compara la calidad del componente XBlock empleando 4 dimensiones de evaluación para herramientas de LA, las cuáles se basan en la metodología GCM. Las dimensiones están relacionadas con los datos, la concientización, reflexión y el impacto. Para la evaluación de las visualizaciones se escogieron estudiantes y maestros de un curso de postgrado de la Universidad de Cuenca. Los participantes tuvieron acceso completo a todas las visualizaciones tomando en cuenta su rol dentro de la plataforma. Adicionalmente, se realizó otra evaluación acerca de la experiencia de usuario aplicando el cuestionario UEQ al XBlock desarrollado en este trabajo de titulación.

Luego de evaluar a los participantes, en sus respuestas se pudo evidenciar el interés por una herramienta que brinde soporte de LA a los estudiantes de un curso. Cerca del 80% de los participantes perciben a XLEA como un estímulo positivo para redirigir su comportamiento y estrategias de aprendizaje o enseñanza hacia el éxito académico dentro del curso. Sin embargo, existen participantes que aseguran que el dashboard puede mejorar su experiencia de usuario añadiendo descripciones y elementos que guíen a aquellos que no han tenido experiencia previa en el manejo de dashboards de LA. En la retroalimentación brindada por los participantes, un 20% expresa que los gráficos deberían cambiar ciertos detalles como colores, tamaños y tipos de letras, tales observaciones están relacionadas con la experiencia de usuario, mientras que, cerca del 80% están de acuerdo en que el dashboard cumple con su objetivo y les permite conocer acerca de su estado actual de aprendizaje o enseñanza con el fin de modificar sus estrategias y comportamiento para alcanzar el éxito académico.

A pesar de esto, existe un alto grado de aceptación y conformidad por parte de los participantes al usar XLEA. La mayoría se enfoca en aquellas visualizaciones que permiten tener una comprensión del comportamiento histórico por semanas de las diferentes actividades que desempeñan dentro del curso, por ejemplo, videos, lecturas y problemas.

Enfocándonos en los resultados de EFLA aplicado a estudiantes, las cuatro dimensiones de evaluación mantienen un puntaje por arriba de los 77/100 puntos, esto demuestra que la herramienta XLEA presenta de manera sencilla y clara los datos en sus visualizaciones, de igual manera, los estudiantes se sienten motivados a reflexionar acerca de su comportamiento y proceso de aprendizaje. De igual manera los resultados del EFLA aplicado a docentes, se observa que las cuatro dimensiones de evaluación mantienen valores por encima de los 77/100 puntos, destacando con 80 puntos la dimensión de datos. Este resultado nos indica que los docentes entienden de qué manera se recopilan y muestran los datos acerca de las interacciones de los estudiantes, permitiéndoles reflexionar acerca de su método de enseñanza y en cómo pueden mejorar en el proceso.

Finalmente, en esta tesis se ha demostrado que es posible crear herramientas de soporte de LA dentro de la plataforma Open edX que permitan tanto a estudiantes y profesores mejorar su proceso de aprendizaje y enseñanza dentro de un curso. Adicionalmente, se ha puesto en evidencia la necesidad de los estudiantes por dashboards como XLEA que les ofrezcan visualizaciones sobre su comportamiento de aprendizaje que antes solo estaban disponibles para los docentes. De esta manera, gracias a las LA se pueden encontrar nuevas formas de innovar el aprendizaje en línea a través de los MOOC y explotar un potencial que antes no se estaba aprovechando.

7.3. Recomendaciones

Luego de haber realizado las evaluaciones, algunos participantes expresaron sus inquietudes y recomendaciones acerca del uso y diseño del dashboard XLEA. La retroalimentación obtenida permitirá crear espacios de mejora en futuras versiones de XLEA o trabajos futuros que deriven del mismo. Se recomienda lo siguiente:

- Debería incluirse una descripción de las funcionalidades de la herramienta, de manera que quede claro lo que se puede obtener de la misma.
- Se podría añadir más información para contextualizar un escenario donde se explique los recursos accedidos por los estudiantes no sólo son videos e información, sino también cuando el estudiante aprueba o reprueba un examen. De esta forma el docente puede percibir los resultados como una retroalimentación de qué temas se puede reforzar para mejorar el proceso de enseñanza-aprendizaje.
- Con respecto a pequeños detalles de la interfaz, la tabla de estadísticas debería indicar claramente las unidades de las columnas, por ejemplo,

cantidad, tiempo, etc. De igual manera, se debería poder acceder a la información de semanas previas.

- Pequeños cambios en los tipos de letras, tamaños, colores, cabeceras de las tablas, que podrían mejorar la experiencia de usuario.

7.4. Trabajos Futuros

Este trabajo de titulación enfocado en el desarrollo de un dashboard de LA y visualizaciones para cursos dentro de la plataforma Open edX, abre un camino para otras investigaciones relacionadas al desarrollo de estos componentes que permitan incluir a los estudiantes en su proceso de aprendizaje.

El alcance del estudio está limitado al análisis del comportamiento de los estudiantes a través de visualizaciones que incorporen indicadores de éxito estudiantil en cursos de la plataforma Open edX. Los resultados obtenidos en este estudio pueden (i) abrir nuevas líneas de investigación centrándose en el mejoramiento y rediseño de material educativo en cursos ofertados en esta plataforma.

Otra línea de investigación podría centrarse en el (ii) procesamiento de los logs de eventos que generan los estudiantes en la plataforma Open edX, ya que actualmente resulta complicado procesar esta información debido a la falta de una estructura concisa y estable de los datos en las diferentes versiones que maneja la plataforma.

Finalmente, existe margen de mejora relacionado al (iii) diseño de visualizaciones de LA dentro de los MOOC, el cual es un área sujeta a cambios e innovaciones.

Referencias

- Al-Shabandar, R., Hussain, A., Laws, A., Keight, R., Lunn, J., & Radi, N. (2017). Machine learning approaches to predict learning outcomes in Massive open online courses. *Proceedings of the International Joint Conference on Neural Networks*, 2017-May, 713–720. <https://doi.org/10.1109/IJCNN.2017.7965922>
- Aldiab, A., Chowdhury, H., Kootsookos, A., Alam, F., & Allhibi, H. (2019). Utilization of Learning Management Systems (LMSs) in higher education system: A case review for Saudi Arabia. *Energy Procedia*, 160, 731–737. <https://doi.org/10.1016/j.egypro.2019.02.186>
- Angel Robledano. (2019a, September 24). *Qué es MySQL: Características y ventajas*. <https://openwebinars.net/blog/que-es-mysql/>
- Angel Robledano. (2019b, October 28). *Qué es MongoDB y características*. <https://openwebinars.net/blog/que-es-mongodb/>
- Antonaci, A., Klemke, R., Kreijns, K., & Specht, M. (2018). Get Gamification of MOOC right! *International Journal of Serious Games*, 5(3), 61–78. <https://doi.org/10.17083/ijsg.v5i3.255>
- Atiaja, L. A., Atiaja, L. A., & Proenza, R. (2016). The MOOCs: origin, characterization, principal problems and challenges in... *Journal of E-Learning and Knowledge Society*, 12(1).
- Audrey Watters. (2013, April 28). *MOOC Mania: Debunking the hype around massive open online courses* | *Library Journal*. <https://www.libraryjournal.com/?detailStory=mooc-mania-debunking-the-hype-around-massive-open-online-courses>
- Ben Kei, D. (2016). Overview of big data and analytics in higher education. In *Big Data and Learning Analytics in Higher Education: Current Theory and Practice* (pp. 1–4). Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_1
- Bennett, L., & Folley, S. (2020). Four design principles for learner dashboards that support student agency and empowerment. *Journal of Applied Research in Higher Education*, 12(1), 15–26. <https://doi.org/10.1108/JARHE-11-2018-0251/FULL/XML>
- Blackwell, A. (2021). *Koa Native Open edX platform Ubuntu 20.04 64 bit*

Installation - Open edX Operations - Confluence.
<https://openedx.atlassian.net/wiki/spaces/OpenOPS/pages/1969455764/Koa+Native+Open+edX+platform+Ubuntu+20.04+64+bit+Installation>

- Boroujeni, M. S., & Dillenbourg, P. (2018). Discovery and temporal analysis of latent study patterns in MOOC interaction sequences. *ACM International Conference Proceeding Series*, 206–215. <https://doi.org/10.1145/3170358.3170388>
- Bote-Lorenzo, M. L., & Gómez-Sánchez, E. (2017). Predicting the decrease of engagement indicators in a MOOC. *ACM International Conference Proceeding Series*, 143–147. <https://doi.org/10.1145/3027385.3027387>
- Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying Learning in the Worldwide Classroom Research into edX's First MOOC. *ERIC*, 8, 13–25. <https://eric.ed.gov/?id=ej1062850>
- Brinton, C. G., Buccapatnam, S., Chiang, M., & Vincent Poor, H. (2015). *Mining MOOC Clickstreams: On the Relationship Between Learner Behavior and Performance*. <https://doi.org/10.48550/arxiv.1503.06489>
- Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). Creating effective learning analytics dashboards: Lessons learnt. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9891 LNCS, 42–56. https://doi.org/10.1007/978-3-319-45153-4_4
- Clow, D. (2013). MOOCs and the funnel of participation. *ACM International Conference Proceeding Series*, 185–189. <https://doi.org/10.1145/2460296.2460332>
- Cobos, R., Gil, S., Lareo, Á., & Vargas, F. A. (2016). Open-DLAs: An open dashboard for learning analytics. *L@S 2016 - Proceedings of the 3rd 2016 ACM Conference on Learning at Scale*, 265–268. <https://doi.org/10.1145/2876034.2893430>
- Coffrin, C., Corrin, L., De Barba, P., & Kennedy, G. (2014). Visualizing patterns of student engagement and performance in MOOCs. *ACM International Conference Proceeding Series*, 83–92. <https://doi.org/10.1145/2567574.2567586>
- Conijn, R., Van den Beemt, A., & Cuijpers, P. (2018). Predicting student performance in a blended MOOC. *Journal of Computer Assisted Learning*, 34(5), 615–628. <https://doi.org/10.1111/JCAL.12270>

- Council, N. R. (2000). *How People Learn*. National Academies Press. <https://doi.org/10.17226/9853>
- Cross, J. S., Keerativoranan, N., Carlon, M. K. J., Tan, Y. H., Rakhimberdina, Z., & Mori, H. (2019). Improving MOOC quality using learning analytics and tools. *Proceedings of 2019 IEEE Learning With MOOCS, LWMOOCS 2019*, 174–179. <https://doi.org/10.1109/LWMOOCS47620.2019.8939617>
- Dacic, D., Sladojevic, S., Lolic, T., & Stefanovic, D. (2019). Process mining possibilities and challenges: A case study. *SISY 2019 - IEEE 17th International Symposium on Intelligent Systems and Informatics, Proceedings*, 161–166. <https://doi.org/10.1109/SISY47553.2019.9111591>
- Dan Davis, Guanliang Chen, Claudia Hauff, & Geert-Jan Houben. (2016). *Gauging MOOC Learners' Adherence to the Designed Learning Path*. <https://eric.ed.gov/?id=ED592664>
- Daphne Koller, Andrew Ng, & Zhenghao Chen. (2013, June 3). *Retention and Intention in Massive Open Online Courses: In Depth*. Educause Review. <https://er.educause.edu/articles/2013/6/retention-and-intention-in-massive-open-online-courses-in-depth>
- de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning*, 32(3), 218–231. <https://doi.org/10.1111/JCAL.12130>
- de Barba, Paula G., Malekian, D., Oliveira, E. A., Bailey, J., Ryan, T., & Kennedy, G. (2020). The importance and meaning of session behaviour in a MOOC. *Computers and Education*, 146, 103772. <https://doi.org/10.1016/j.compedu.2019.103772>
- DeBoer, J., Ho, A. D., Stump, G. S., & Breslow, L. (2014). Changing “Course”: Reconceptualizing Educational Variables for Massive Open Online Courses. *Educational Researcher*, 43(2), 74–84. <https://doi.org/10.3102/0013189X14523038>
- Deploy Your XBlock in Devstack — Open edX XBlock Tutorial documentation*. (n.d.). Retrieved April 2, 2022, from https://edx.readthedocs.io/projects/xblock-tutorial/en/latest/edx_platform/devstack.html#enable-the-xblock-in-your-course
- Dhawal Shah. (2020, November 30). *MOOCs in 2020*.

<https://www.classcentral.com/report/mooc-stats-2020/>

edutechnica. (2018). *6th Annual LMS Data Update*.
<https://edutechnica.com/2018/10/06/6th-annual-lms-data-update/>

edX. (2016). *EdX Research Guide*.
https://edx.readthedocs.io/projects/devdata/en/stable/internal_data_formats/tracking_logs.html

edX. (2018). *Overview of EdX Insights*. <https://edx.readthedocs.io/projects/edx-insights/en/latest/Overview.html>

edX. (2021). *EdX Open Learning XML Guide*.
<https://edx.readthedocs.io/projects/edx-open-learning-xml/en/latest/index.html#>

edX. (2022, March 5). *Watching Videos on the edX Video Player*.
https://edx.readthedocs.io/projects/open-edx-learner-guide/en/latest/SFD_video_player.html#video-player

edX Inc. (2021). *Introduction to XBlocks*.

edX Insights. (n.d.). Retrieved March 27, 2022, from <https://insights.edx.org/>

Et-Taleby, A., Boussetta, M., & Benslimane, M. (2020). Faults detection for photovoltaic field based on k-means, elbow, and average silhouette techniques through the segmentation of a thermal image. *International Journal of Photoenergy*, 2020. <https://doi.org/10.1155/2020/6617597>

Fernández Buenaño, D., & Luján-Mora, S. (2019). RUA: Propuesta de aplicación de minería de procesos para evaluar las rutas de aprendizaje de estudiantes con discapacidad visual en cursos en línea. *RISTI*, N.º E17, 01/2019, 1035–1047. <http://rua.ua.es/dspace/handle/10045/88328>

Fluxicon. (2022). *Process Mining and Automated Process Discovery Software for Professionals*. <https://fluxicon.com/discover/>

Gaebel, M. (2013). *MOOCs Massive Open Online Courses*.
<https://www.eua.eu/downloads/publications/moocs-massive-open-online-courses.pdf>

Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs. *User Modeling and User-Adapted Interaction*, 28(2), 127–203.
<https://doi.org/10.1007/s11257-018-9203-z>

- Gong, F., Bu, F., Zhang, Y., Yan, Y., Hu, R., & Dong, M. (2020). Visual Clustering Analysis of Electricity Data Based on t-SNE. *2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics, ICCCBDA 2020*, 234–240. <https://doi.org/10.1109/ICCCBDA49378.2020.9095574>
- Hilliger, I., De Laet, T., Henríquez, V., Guerra, J., Ortiz-Rojas, M., Zuñiga, M. Á., Baier, J., & Pérez-Sanagustín, M. (2020). For Learners, with Learners: Identifying Indicators for an Academic Advising Dashboard for Students. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12315 LNCS, 117–130. https://doi.org/10.1007/978-3-030-57717-9_9
- Ioana, J. (2021). *The Dashboard That Loved Me: Designing adaptive learning analytics for self-regulated learning — Open Universiteit research portal* [Open Universiteit]. <https://research.ou.nl/en/publications/the-dashboard-that-loved-me-designing-adaptive-learning-analytics>
- IONOS. (2019, September 25). *SDK: ¿qué es el software development kit?* <https://www.ionos.es/digitalguide/paginas-web/desarrollo-web/software-development-kit/>
- Jan Štrobl, Marek Piorecký, & Vladimír Krajča. (2017). *Methods for automatic estimation of the number of clusters for k-means algorithm used on eeg signal: feasibility study*. 47, 81–87. <https://ojs.cvut.cz/ojs/index.php/CTJ/article/view/4474>
- Javier, H., Díaz, P., & Merino, P. M. (2015). *Learning Analytics Visualizations of Student-Activity Time Distribution for the Open Edx Platform*. <https://e-archivo.uc3m.es/handle/10016/26077>
- Jim, W. (2012). *What You Need to Know About MOOCs*. <https://www.chronicle.com/article/what-you-need-to-know-about-moocs/>
- Kane, & Trochim, W. M. K. (2009). Concept Mapping for Planning and Evaluation. *Http://Dx.Doi.Org/10.1177/1558689808326121*, 3(1), 87–89. <https://doi.org/10.1177/1558689808326121>
- Kennedy, G., De Barba, P., Coffrin, C., & Corrin, L. (2015). Predicting success: How learners' prior knowledge, skills and activities predict MOOC performance. *ACM International Conference Proceeding Series, 16-20-Marc*, 136–140. <https://doi.org/10.1145/2723576.2723593>
- Khalil, M., & Belokry, G. (2020). OXALIC: An Open edX Advanced Learning Analytics Tool. *Proceedings of 2020 IEEE Learning With MOOCs*,

- Khalil, M., Khalil, M., & Ebner, M. (2015). Learning Analytics: Principles and Constraints. *EdMedia + Innovate Learning*, 2015(1), 1789–1799.
- Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *ACM International Conference Proceeding Series*, 170–179. <https://doi.org/10.1145/2460296.2460330>
- Kloft, M., Stiehler, F., Zheng, Z., & Pinkwart, N. (2014). *Predicting MOOC Dropout over Weeks Using Machine Learning Methods*.
- Koller, D., Ng, A., B. Do, C., & Chen, Z. (2013). Retention and intention in massive open online courses: In depth. *Educause Review*, 48(3), 62–63.
- Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). Penetrating the black box of time-on-task estimation. *ACM International Conference Proceeding Series, 16-20-March-2015*, 184–193. <https://doi.org/10.1145/2723576.2723623>
- Kovanović, V., Michail, C., Joksimović, S., Siemens, G., Katerinopoulos, P., & Gašević, D. (2017). Developing a MOOC experimentation platform: Insights from a user study. *ACM International Conference Proceeding Series*, 1–5. <https://doi.org/10.1145/3027385.3027398>
- Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5298 LNCS, 63–76. https://doi.org/10.1007/978-3-540-89350-9_6
- Liang, J., Li, C., & Zheng, L. (2016). Machine learning application in MOOCs: Dropout prediction. *ICCSE 2016 - 11th International Conference on Computer Science and Education*, 52–57. <https://doi.org/10.1109/ICCSE.2016.7581554>
- Liang, J., Yang, J., Wu, Y., Li, C., & Zheng, L. (2016). Big data application in education: Dropout prediction in edx MOOCs. *Proceedings - 2016 IEEE 2nd International Conference on Multimedia Big Data, BigMM 2016*, 440–443. <https://doi.org/10.1109/BIGMM.2016.70>
- Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *International Review*

of *Research in Open and Distance Learning*, 14(3), 202–227.
<https://doi.org/10.19173/irrodl.v14i3.1455>

- Loeckx, J. (2016). Blurring boundaries in education: Context and impact of MOOCs. *International Review of Research in Open and Distance Learning*, 17(3), 92–121. <https://doi.org/10.19173/irrodl.v17i3.2395>
- Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Muñoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, 80, 179–196. <https://doi.org/10.1016/j.chb.2017.11.011>
- Maldonado-Mahauad, J., Pérez-Sanagustín, M., Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., & Delgado-Kloos, C. (2018). Predicting Learners' Success in a Self-paced MOOC Through Sequence Patterns of Self-regulated Learning. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11082 LNCS, 355–369. https://doi.org/10.1007/978-3-319-98572-5_27
- Maldonado, J. J., Palta, R., Vazquez, J., Bermeo, J. L., Perez-Sanagustin, M., & Muñoz-Gama, J. (2016). Exploring differences in how learners navigate in MOOCs based on self-regulated learning and learning styles: A process mining approach. *Proceedings of the 2016 42nd Latin American Computing Conference, CLEI 2016*. <https://doi.org/10.1109/CLEI.2016.7833356>
- Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2016). LATUX: an Iterative Workflow for Designing, Validating and Deploying Learning Analytics Visualisations. *Journal of Learning Analytics*, 2(3), 9–39. <https://doi.org/10.18608/jla.2015.23.3>
- McAuley, A., Bonnie, S., Siemens, G., & Cormier, D. (2010). *The MOOC model for digital practice*. https://www.academia.edu/22561210/The_MOOC_model_for_digital_practice?auto=citations&from=cover_page
- McKinney, W. (2018). Python for Data Analysis : Data Wrangling with Pandas, NumPy, and IPython. *Transplantation*, 71(10), 1385–1389.
- Mukala, P., Buijs, & van der Aalst, W. (2015). *Exploring students' learning behaviour in MOOCs using process mining techniques — Eindhoven University of Technology research portal*. <https://research.tue.nl/en/publications/exploring-students-learning->

behaviour-in-moocs-using-process-mini

Nagrecha, S., Dillon, J. Z., & Chawla, N. V. (2017). MOOC dropout prediction: Lessons learned from making pipelines interpretable. *26th International World Wide Web Conference 2017, WWW 2017 Companion*, 351–359. <https://doi.org/10.1145/3041021.3054162>

Nemoto, T., & Beglar, D. (2014). c. *JALT2013 Conference Proceedings*, 1–8. https://dl.ndl.go.jp/info:ndljp/pid/8815334?itemId=info%3Andljp%2Fpid%2F8815334&__lang=en

Open edX. (n.d.). 6.3. *User Info and Learner Progress Data — EdX Research Guide documentation*. Retrieved March 7, 2022, from https://edx.readthedocs.io/projects/devdata/en/latest/internal_data_formats/sql_schema.html

Open edX. (2022a). *Getting Started with the XBlock SDK*. https://edx.readthedocs.io/projects/xblock-tutorial/en/latest/sdk/get_started_sdk.html

Open edX. (2022b). *Open edX Architecture*.

Pardos, Z. A., Whyte, A., & Kao, K. (2016). moocRP: Enabling Open Learning Analytics with an Open Source Platform for Data Distribution, Analysis, and Visualization. *Technology, Knowledge and Learning*, 21(1), 75–98. <https://doi.org/10.1007/s10758-015-9268-2>

Patterns of engagement in connectivist MOOCs - Enlighten: Publications. (n.d.). Retrieved May 11, 2022, from <https://eprints.gla.ac.uk/186192/>

Perez-Alvarez, R. A., Maldonado-Mahauad, J., Sharma, K., Sapunar-Opazo, D., & Perez-Sanagustin, M. (2020). Characterizing Learners' Engagement in MOOCs: An Observational Case Study Using the NoteMyProgress Tool for Supporting Self-Regulation. *IEEE Transactions on Learning Technologies*, 13(4), 676–688. <https://doi.org/10.1109/TLT.2020.3003220>

Pijera Díaz, H. J., Ruiz, J. S., Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., & Delgado Kloos, C. (2015). Using video visualizations in open edX to understand learning interactions of students. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9307, 522–525. https://doi.org/10.1007/978-3-319-24258-3_51

Purwoningsih, T., Santoso, H. B., & Hasibuan, Z. A. (2019a, October 1). Online Learners' Behaviors Detection Using Exploratory Data Analysis and Machine

- Learning Approach. *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*.
<https://doi.org/10.1109/ICIC47613.2019.8985918>
- Purwoningsih, T., Santoso, H. B., & Hasibuan, Z. A. (2019b). Online Learners' Behaviors Detection Using Exploratory Data Analysis and Machine Learning Approach. *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*.
<https://doi.org/10.1109/ICIC47613.2019.8985918>
- Ramesh, A., Goldwasser, D., Huang, B., Daume, H., & Getoor, L. (2014). Learning latent engagement patterns of students in online courses | Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. *ACM Digital Library*, 1272–1278.
<https://dl.acm.org/doi/abs/10.5555/2893873.2894071>
- Rohloff, T., Sauer, D., & Meinel, C. (2019, December 1). Student Perception of a Learner Dashboard in MOOCs to Encourage Self-Regulated Learning. *TALE 2019 - 2019 IEEE International Conference on Engineering, Technology and Education*. <https://doi.org/10.1109/TALE48000.2019.9225939>
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Delgado Kloos, C. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior*, 47, 139–148. <https://doi.org/10.1016/j.chb.2014.07.002>
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Pijeira Díaz, H. J., Ruiz, J. S., & Kloos, C. D. (2017). Evaluation of a learning analytics application for open edX platform. *Computer Science and Information Systems*, 14(1), 51–73. <https://doi.org/10.2298/CSIS160331043R>
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Pijeira Díaz, H. J., Santofimia Ruiz, J., & Delgado Kloos, C. (2017). *Evaluation of a Learning Analytics Application for Open edX Platform*. <https://doi.org/10.2298/CSIS160331043R>
- Ruiz, J. S., Díaz, H. J. P., Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., & Kloos, C. D. (2014). Towards the development of a learning analytics extension in open edX. *ACM International Conference Proceeding Series*, 299–306. <https://doi.org/10.1145/2669711.2669914>
- Samuelson, J., & Khalil, M. (2020). Study Effort and Student Success: A MOOC Case Study. *Advances in Intelligent Systems and Computing*, 916, 215–226. https://doi.org/10.1007/978-3-030-11932-4_22

- Santofimia Ruiz, J. (2017). *Diseño de arquitectura y visualizaciones para un módulo de analítica del aprendizaje en Open edX*. <https://e-archivo.uc3m.es/handle/10016/26041?show=full>
- Scheffel, M., Niemann, K., & Jivet, I. (2017). *The Evaluation Framework for Learning Analytics*.
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6), 103. <https://doi.org/10.9781/IJIMAI.2017.09.001>
- Schunk, D. H., & Zimmerman, B. J. (2012). Motivation and self-regulated learning: Theory, research, and applications. *Motivation and Self-Regulated Learning: Theory, Research, and Applications*, 1–417. <https://doi.org/10.4324/9780203831076>
- Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D. E. (2014). Who does what in a massive open online course? *Communications of the ACM*, 57(4), 58–65. <https://doi.org/10.1145/2500876>
- Sedrakyan, G., Mannens, E., & Verbert, K. (2019). Guiding the choice of learning dashboard visualizations: Linking dashboard design and data visualization concepts. *Journal of Computer Languages*, 50, 19–38. <https://doi.org/10.1016/J.JVLC.2018.11.002>
- Shah, D. (2020a). *By The Numbers: MOOCs in 2020 — Class Central*. <https://www.classcentral.com/report/mooc-stats-2020/>
- Shah, D. (2020b). *By The Numbers: MOOCs in 2020 — Class Central*.
- Shah, D. (2020c). *The Second Year of The MOOC: A Review of MOOC Stats and Trends in 2020 — Class Central*. <https://www.classcentral.com/report/the-second-year-of-the-mooc/>
- Shi, Conglei, Fu, S., Chen, Q., & Qu, H. (2015). VisMOOC: Visualizing video clickstream data from Massive Open Online Courses. *IEEE Pacific Visualization Symposium, 2015-July*, 159–166. <https://doi.org/10.1109/PACIFICVIS.2015.7156373>
- Shi, Congming, Wei, B., Wei, S., Wang, W., Liu, H., & Liu, J. (2021). A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *Eurasip Journal on Wireless Communications and Networking*, 2021(1), 1–16. <https://doi.org/10.1186/S13638-021-01910-W/FIGURES/6>

- Sinha, T., Jermann, P., Li, N., & Dillenbourg, P. (2014). *Your click decides your fate: Inferring Information Processing and Attrition Behavior from MOOC Video Clickstream Interactions*. 3–14. <https://doi.org/10.48550/arxiv.1407.7131>
- Smith, B., & Eng, M. (2013). MOOCs: A learning journey - Two continuing education practitioners investigate and compare cMOOC and xMOOC learning models and experiences. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8038 LNCS, 244–255. https://doi.org/10.1007/978-3-642-39750-9_23
- Taylor, C., Veeramachaneni, K., & O'Reilly, U.-M. (2014). *Likely to stop? Predicting Stopout in Massive Open Online Courses*. <http://arxiv.org/abs/1408.3382>
- Torre, M. V., Tan, E., & Hauff, C. (2020). EdX log data analysis made easy: Introducing ELAT: An open-source, privacy-aware and browser-based edX log data analysis tool. *ACM International Conference Proceeding Series*, 502–511. <https://doi.org/10.1145/3375462.3375510>
- Trafford, P., & Shirota, Y. (2011). An Introduction to Virtual Learning Environments. *Gakushuin Economic Papers*, 48(3), 143–151. <https://econpapers.repec.org/RePEc:abc:gakuep:48-3-1>
- UBUCEV. (2019). *MOOC ¿Qué son?* Centro de Enseñanza Virtual de La Universidad de Burgos. <https://www3.ubu.es/ubucevblog/mooc-que-son/>
- van den Beemt, A., Buys, J., & van der Aalst, W. (2018a). Analysing structured learning behaviour in Massive Open Online Courses (MOOCs): An approach based on process mining and clustering. *International Review of Research in Open and Distance Learning*, 19(5), 38–60. <https://doi.org/10.19173/irrodl.v19i5.3748>
- van den Beemt, A., Buys, J., & van der Aalst, W. (2018b). Analysing structured learning behaviour in Massive Open Online Courses (MOOCs): An approach based on process mining and clustering. *International Review of Research in Open and Distance Learning*, 19(5), 38–60. <https://doi.org/10.19173/irrodl.v19i5.3748>
- Van Der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, 16(9), 1128–1142. <https://doi.org/10.1109/TKDE.2004.47>

- Van Eck, M. L., Lu, X., Leemans, S. J. J., & Van Der Aalst, W. M. P. (2015). A Process Mining Project Methodology. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9097, 297–313. https://doi.org/10.1007/978-3-319-19069-3_19
- Vermunt, J. D. (2005). Relations between student learning patterns and personal and contextual factors and academic performance. *Higher Education*, 49(3), 205–234. <https://doi.org/10.1007/s10734-004-6664-2>
- Vermunt, J. D., & Donche, V. (2017). A Learning Patterns Perspective on Student Learning in Higher Education: State of the Art and Moving Forward. *Educational Psychology Review*, 29(2), 269–299. <https://doi.org/10.1007/s10648-017-9414-6>
- Volchek, D., Romanov, A., & Mouromtsev, D. (2017). Towards the semantic MOOC: Extracting, enriching and interlinking E-learning data in open edx platform. *Communications in Computer and Information Science*, 786, 295–305. https://doi.org/10.1007/978-3-319-69548-8_20
- Watson, S. L., Loizzo, J., Watson, W. R., Mueller, C., Lim, J., & Ertmer, P. A. (2016). Instructional design, facilitation, and perceived learning outcomes: an exploratory case study of a human trafficking MOOC for attitudinal change. *Educational Technology Research and Development*, 64(6), 1273–1300. <https://doi.org/10.1007/s11423-016-9457-2>
- West, D., Luzecky, A., Toohey, D., Vanderlelie, J., & Searle, B. (2020). Do academics and university administrators really know better? The ethics of positioning student perspectives in learning analytics. *Australasian Journal of Educational Technology*, 36(2), 60–70. <https://doi.org/10.14742/AJET.4653>
- Williams, K. M., Stafford, R. E., Corliss, S. B., & Reilly, E. D. (2018). Examining student characteristics, goals, and engagement in Massive Open Online Courses. *Computers and Education*, 126, 433–442. <https://doi.org/10.1016/j.compedu.2018.08.014>
- XBlock Installation in Open edX Lilac - Open edX discussions*. (n.d.). Retrieved March 31, 2022, from <https://discuss.openedx.org/t/xblock-installation-in-open-edx-lilac-edx-bin-pip-edxapp-command-not-found-message/4940/4>
- Zhang, X. Da. (2020). A matrix algebra approach to artificial intelligence. *A Matrix Algebra Approach to Artificial Intelligence*, 223–240. <https://doi.org/10.1007/978-981-15-2770-8>

Anexos

Anexo 1: Evento guardado en el archivo tracking logs

```
{
  "agent": "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like
  Gecko) Chrome/30.0.1599.101 Safari/537.36",
  "context": {
    "course_id": "edx/AN101/2014_T1",
    "module": {
      "display_name": "Multiple Choice Questions"
    },
    "org_id": "edx",
    "user_id": 9999999
  },
  "event": {
    "answers": {
      "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_2_1":
      "yellow",
      "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_4_1": [
        "choice_0",
        "choice_2"
      ]
    },
    "attempts": 1,
    "correct_map": {
      "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_2_1": {
        "correctness": "incorrect",
        "hint": "",
        "hintmode": null,
        "msg": "",
        "npoints": null,
        "queuestate": null
      },
      "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_4_1": {
        "correctness": "correct",
        "hint": "",
        "hintmode": null,
        "msg": "",
        "npoints": null,
        "queuestate": null
      }
    },
    "grade": 2,
    "max_grade": 3,
    "problem_id":
    "i4x://edx/AN101/problem/a0effb954cca4759994f1ac9e9434bf4",
    "state": {
      "correct_map": {},
      "done": null,
      "input_state": {
        "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_2_1":
        {},
        "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_4_1":
        {}
      }
    }
  },
}
```

UCUENCA

```
    "seed": 1,
    "student_answers": {}
  },
  "submission": {
    "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_2_1": {
      "answer": "yellow",
      "correct": false,
      "input_type": "optioninput",
      "question": "What color is the open ocean on a sunny day?",
      "response_type": "optionresponse",
      "variant": ""
    },
    "i4x-edx-AN101-problem-a0effb954cca4759994f1ac9e9434bf4_4_1": {
      "answer": [
        "a piano",
        "a guitar"
      ],
      "correct": true,
      "input_type": "checkboxgroup",
      "question": "Which of the following are musical
instruments?",
      "response_type": "choiceresponse",
      "variant": ""
    }
  },
  "success": "incorrect"
},
"event_source": "server",
"event_type": "problem_check",
"host": "precise64",
"referer":
"http://localhost:8001/container/i4x://edX/DemoX/vertical/69dedd38233
a46fc89e4d7b5e8dalbf4?action=new",
"accept_language": "en-US,en;q=0.8",
"ip": "NN.N.N.N",
"page": "x_module",
"time": "2014-03-03T16:19:05.584523+00:00",
"username": "AAAAAAAAAA"
}
```

Anexo 2: Descripción de las tablas que contienen información acerca de los estudiantes dentro de la plataforma.

Tabla	Descripción
auth_user	La tabla está integrada en el framework web edX Django. Contiene información necesaria para la entrada de usuario y permisos.
auth_userprofile	La tabla almacena los datos demográficos de los usuarios recopilados cuando los alumnos se registran para obtener una cuenta de usuario o agregan información

Miguel Ángel Macías Narváez
Jonathan Henry Campoberde Ávila

UCUENCA

	de perfil sobre sí mismos. Cada fila de esta tabla corresponde a una fila en auth_user
certificates_generatedcertificate	La tabla rastrea el estado de los certificados que se han emitido para un curso. Puede utilizar esta tabla para saber cuál de sus alumnos recibió un certificado.
completion_blockcompletion	Recursos que han completado los usuarios
course_creators_coursecreator	Los usuarios que han creado los cursos
course_overviews_courseoverview	Se tiene las fechas de inicio y de fin de un curso, pequeña descripción del curso, la nota mínima para pasar y el nombre del curso
course_overviews_courseoverviewtab	Pestañas de cada curso como home,progress y discussion de cada curso
course_overviews_historicalcourseoverview	Tiene un historial de los cambios que se han realizado en el curso en las notas
courseware_studentmodule	La tabla contiene el estado más actual del curso, incluido el envío del problema más reciente y la unidad visitada en cada subsección.
dark_lang_darklangconfig	Registro de las personas que han cambiado el lenguaje
django_migrations	Un historial de las migraciones que se han realizado
embargo_country	Código país de los estudiantes
grades_persistentcoursegrade	La tabla almacena valores persistentes para las calificaciones de los cursos de los alumnos.
grades_persistentsubsectiongrade	La tabla almacena valores persistentes para las calificaciones de subsección de los alumnos.
student_anonymoususerid	Contiene el id anónimo de un estudiante en un curso

student_courseaccessrole	El tipo de acceso que se tiene de un usuario en específico
student_courseenrollment	Cada fila en esta tabla representa la inscripción de un alumno en algún curso en particular. Se crea una fila para cada alumno que comienza el proceso de inscripción, incluso si nunca completa el registro en la plataforma activando la cuenta de usuario.
user_api_usercoursetag	Esta tabla utiliza pares clave-valor para almacenar metadatos sobre la participación de un alumno específico en un curso específico. Por ejemplo, para un curso que asigna alumnos a grupos de forma aleatoria para experimentos de contenido, una fila en esta tabla identifica la asignación del alumno a una partición y un grupo.
user_api_userpreference	Preferencias de usuario. ej. idioma
user_tasks_usertaskstatus	Tareas que realiza el estudiante en la plataforma. ej. exportar/importar cursos
verify_student_verificationstatus	La tabla mostraba los intentos y los resultados de la re-verificación del alumno.
verify_student_verificationdeadline	Especifica la fecha cuando termina el curso

Anexo 3: Scripts y comandos para migrar los datos

a) Script para migrar los datos de Ironwood

```
#!/bin/bash
MYSQL_CONN="-uroot -p"
echo "Reading MySQL database names..."
mysql ${MYSQL_CONN} -ANe "SELECT schema_name FROM information_schema.schemata
WHERE schema_name NOT IN ('mysql','information_schema','performance_schema',
'sys')" > /tmp/db.txt
DBS="--databases $(cat /tmp/db.txt)"
NOW="$(date +%Y%m%dT%H%M%S)"
SQL_FILE="mysql-data-`${NOW}`.sql"
echo "Dumping MySQL structures..."
mysqldump ${MYSQL_CONN} --add-drop-database --skip-add-drop-table --no-data
`${DBS}` > `${SQL_FILE}`
echo "Dumping MySQL data..."
# If there is table data you don't need, add --ignore-table=tablename
```

UCUENCA

```
mysqldump ${MYSQL_CONN} --no-create-info ${DBS} >> ${SQL_FILE}

for db in edxapp cs_comments_service; do
    echo "Dumping Mongo db ${db}..."
    mongodump -u admin -p -h localhost --authenticationDatabase admin -d ${db}
    --out mongo-dump-${NOW}
done

tar -czf openedx-data-${NOW}.tgz ${SQL_FILE} mongo-dump-${NOW}
```

b) Comandos para restaurar los datos en la máquina de Koa

```
tar -xvf openedx-data-20200411T154750.tgz
mysql -uroot -p < mysql-data-20200411T154750.sql
mongorestore -u admin -p -h localhost --authenticationDatabase admin --drop -d
edxapp mongo-dump-20200411T154750/edxapp
mongorestore -u admin -p -h localhost --authenticationDatabase admin --drop -d
cs_comment_service mongo-dump-20200411T154750/cs_comment_service_development
```

c) Comandos para para actualizar y validar los datos en la instancia de Koa

```
/edx/app/edx_ansible/edx_ansible/util/install/native.sh --tags migrate
```

Anexo 4: Scripts que analiza los eventos generados por los estudiantes

```
def forumData(self, data):
    """[Process data generated in the logs about the forums]
    Args:
        data ([Dataframe]): [Data extracted from the logs]

    Returns:
        [dict]: [Dictionary with forum data processed ]
    """
    try:
        forum_event_types =
('edx.forum.comment.created', 'edx.forum.response.created', 'edx.forum.thread.cr
eated')
        forum_data = data[data["event_type"].isin(forum_event_types)]

        dict_data_forum = []
        for row in forum_data.to_dict('records'):
            dict_data = {}
            dict_data['user_id'] = row['context']['user_id']
            dict_data['course_id'] = row['context']['course_id']
            dict_data['block_id'] = ""
            dict_data['event_type'] = ""
            dict_data['time'] = row['time']
            dict_data['category'] = 'foro'
```

UCUENCA

```
        if ("edx.forum.thread.created" in row['name']):
            dict_data['event_type'] = "Pregunta foros"
        elif ('edx.forum.response.created' in row['name']):
            dict_data['event_type'] = "Responde foro"
        elif('edx.forum.comment.created' in row['name']):
            dict_data['event_type'] = "Comenta post"

        dict_data_forum.append(dict_data)

    return self.saveDataProcessedFile(dict_data_forum)

except Exception as e:
    return False

def videoData(self,data):
    try:
        video_general_event_types =
('load_video','edx.video.loaded','pause_video',
'edx.video.paused','play_video', 'edx.video.played', 'stop_video',
'edx.video.stopped')
        video_data_general =
data[data["event_type"].isin(video_general_event_types)]
        dict_data_video_general = []
        storagedSucessfully = True

        for row in video_data_general.to_dict('records'):
            dict_data = {}
            event_object = loads(row['event'])
            dict_data['user_id'] = row['context']['user_id']
            dict_data['course_id'] = row['context']['course_id']
            dict_data['block_id'] = event_object['id']
            dict_data['event_type'] = row['event_type']
            dict_data['time'] = row['time']
            dict_data['category'] = 'video'
            dict_data_video_general.append(dict_data)

        video_event_positin = ('seek_video', 'edx.video.position.changed')
        video_data_position =
data[data["event_type"].isin(video_event_positin)]

        dict_data_video_position = []
        for row in video_data_position.to_dict('records'):
            dict_data = {}
            event_object = loads(row['event'])
            dict_data['user_id'] = row['context']['user_id']
            dict_data['course_id'] = row['context']['course_id']
            dict_data['block_id'] = event_object['id']
            dict_data['event_type'] = "Adelanta video" if
(event_object["new_time"] > event_object["old_time"]) else "Retrocede video"
            dict_data['time'] = row['time']
            dict_data['category'] = 'video'
            dict_data_video_position.append(dict_data)

        xblock_events = (r'(video\+block)')
        block_visited_video =
data.loc[data["event_type"].str.contains(xblock_events,
case=False,regex=True)]
```

UCUENCA

```
dict_data_video_completed = []
user_start = {}

for row in block_visited_video.to_dict('records'):
    dict_data = {}
    user_id = row['context']['user_id']
    time = row['time']
    course_id = row['context']['course_id']

    if ("publish_completion" in row["event_type"]):
        block_id = (row["context"]["path"].split("@")[-
1]).split("/")[-3]
        if (user_id not in user_start):
            user_start[user_id] = {}
            user_start[user_id][block_id] = "1"
            dict_data['user_id'] = user_id
            dict_data['course_id'] = course_id
            dict_data['block_id'] = block_id
            dict_data['event_type'] = "Completa video"
            dict_data['time'] = time
            dict_data['category'] = 'video'

            dict_data_video_completed.append(dict_data)
        else:
            if (block_id in user_start[user_id]):
                dict_data['user_id'] = user_id
                dict_data['course_id'] = course_id
                dict_data['block_id'] = block_id
                dict_data['event_type'] = "Repite video"
                dict_data['time'] = time
                dict_data['category'] = 'video'

                dict_data_video_completed.append(dict_data)
            else:
                user_start[user_id][block_id] = "1"
                dict_data['user_id'] = user_id
                dict_data['course_id'] = course_id
                dict_data['block_id'] = block_id
                dict_data['event_type'] = "Completa video"
                dict_data['time'] = time
                dict_data['category'] = 'video'
                dict_data_video_completed.append(dict_data)

    return True
except Exception as e:
    return False

def processAssesment(self, data):
    try:
        assessments_event_types = ('problem_check',)
        assessments_data =
data[data["event_type"].isin(assessments_event_types)]
        assessments_data =
assessments_data[assessments_data['event_source']=='server']
        dict_data_problem_check = []
        user_start = {}
```

UCUENCA

```
for row in assessments_data.to_dict('records'):
    dict_data = {}
    user_id = row['context']['user_id']
    time = row['time']
    course_id = row['context']['course_id']
    block_id = (row["event"]["problem_id"].split("@")[-1])

    if (user_id in user_start):
        if (block_id in user_start[user_id]):
            dict_data['user_id'] = user_id
            dict_data['course_id'] = course_id
            dict_data['block_id'] = block_id
            dict_data['event_type'] = "Repite problema"
            dict_data['time'] = time
            dict_data['category'] = 'problema'
            dict_data_problem_check.append(dict_data)
        else:

            user_start[user_id][block_id] = "1"
            dict_data['user_id'] = user_id
            dict_data['course_id'] = course_id
            dict_data['block_id'] = block_id
            dict_data['event_type'] = "Completa problema"
            dict_data['time'] = time
            dict_data['category'] = 'problema'
            dict_data_problem_check.append(dict_data)
    else:
        user_start[user_id] = {}
        user_start[user_id][block_id] = "1"
        dict_data['user_id'] = user_id
        dict_data['course_id'] = course_id
        dict_data['block_id'] = block_id
        dict_data['event_type'] = "Completa problema"
        dict_data['time'] = time
        dict_data['category'] = 'problema'
        dict_data_problem_check.append(dict_data)

    return True
except Exception as e:
    return False

def processLectures(self, data):
    try:
        xblock_events_html = (r'(html\+block)')
        block_html =
data.loc[data["event_type"].str.contains(xblock_events_html,
case=False, regex=True)]
        block_html.reset_index(drop=True, inplace=True)
        dict_data_html = []
        user_start = {}

        for row in block_html.to_dict('records'):
            dict_data = {}
            user_id = row['context']['user_id']
            time = row['time']
            course_id = row['context']['course_id']
            if (course_id==''):
```

UCUENCA

```
        continue

    if ("publish_completion" in row["event_type"]):
        block_id = (row["context"]["path"].split("@")[-1]).split("/")[-3]
        if (user_id not in user_start):
            user_start[user_id] = {}
            user_start[user_id][block_id] = "1"
            dict_data['user_id'] = user_id
            dict_data['course_id'] = course_id
            dict_data['block_id'] = block_id
            dict_data['event_type'] = "Completa lectura"
            dict_data['time'] = time
            dict_data['category'] = 'lectura'
            dict_data_html.append(dict_data)
        else:
            if (block_id in user_start[user_id]):
                dict_data['user_id'] = user_id
                dict_data['course_id'] = course_id
                dict_data['block_id'] = block_id
                dict_data['event_type'] = "Repite lectura"
                dict_data['time'] = time
                dict_data['category'] = 'lectura'
                dict_data_html.append(dict_data)
            else:
                user_start[user_id][block_id] = "1"
                dict_data['user_id'] = user_id
                dict_data['course_id'] = course_id
                dict_data['block_id'] = block_id
                dict_data['event_type'] = "Completa lectura"
                dict_data['time'] = time
                dict_data['category'] = 'lectura'
                dict_data_html.append(dict_data)
        else:
            block_id = (row["context"]["path"].split("@")[-1])
            if (user_id in user_start):#Verify if the user start a
lecture previously

                if (block_id in user_start[user_id]):
                    dict_data['user_id'] = user_id
                    dict_data['course_id'] = course_id
                    dict_data['block_id'] = block_id
                    dict_data['event_type'] = "Repite lectura"
                    dict_data['time'] = time
                    dict_data['category'] = 'lectura'
                    dict_data_html.append(dict_data)
                else:
                    user_start[user_id][block_id] = "1"
                    dict_data['user_id'] = user_id
                    dict_data['course_id'] = course_id
                    dict_data['block_id'] = block_id
                    dict_data['event_type'] = "Completa lectura"
                    dict_data['time'] = time
                    dict_data['category'] = 'lectura'
                    dict_data_html.append(dict_data)
            else:
                pass
    return True
```

```
except Exception as e:  
    StorageLog.save('Error en procesar los datos de la lectura:  
'+str(e))  
    return False
```

Anexo 5: Enlace al repositorio de Github con los Scripts que analiza las actividades realizadas y obtiene estadísticas en forma de sesiones.

https://github.com/JonnHenry/xlea_server.git

Anexo 6: Artículo publicado en la conferencia LACLO 2021 con doi: 10.1109/LACLO54177.2021.00088

Proposal for the Design and Implementation of a XBlock in Open edX to Support Learning Analytics

Jonnathan Campoberde
Departamento de Ciencias de la
Computación
Universidad de Cuenca
Cuenca, Ecuador
jonnathan.campoberde@ucuenca.edu.ec

Miguel A. Macías
Departamento de Ciencias de la
Computación
Universidad de Cuenca
Cuenca, Ecuador
mangel.maciasn@ucuenca.edu.ec

Jorge Maldonado-Mahauad
Departamento de Ciencias de la
Computación
Universidad de Cuenca
Cuenca, Ecuador
jorge.maldonado@ucuenca.edu.ec

Abstract— Nowadays, when users using information systems, they generate a large amount of data leaving behind a trace as a result of their interaction; e.g., when students access educational materials on Virtual Learning Environments (VLE). This has developed into what we now know as Massive Open Online Courses (MOOCs), which have millions of registered students. Due to the large amount of data that is generated within these MOOCs, Learning Analytics (LA) has emerged as an alternative to improve teaching and learning processes through data analysis. Open edX in an attempt to incorporate Learning Analytics into its Insights development platform, which provides very simple visualizations. Likewise, other projects have been added, which over time have become obsolete or do not provide sufficient support to improve the learning process of students. Therefore, this study proposes the design, development and evaluation of a learning analytics dashboard for the Open edX platform. The tool will incorporate indicators of student success in its visualizations in order to improve the learning process within the platform.

Keywords— Learning Analytics, Open edX, dashboard, student success, XBlock

I. INTRODUCCIÓN

En un modelo tradicional de educación el profesor es el encargado de compartir sus enseñanzas y habilidades a sus estudiantes dentro de una institución. Los estudiantes demuestran el conocimiento adquirido por medio de actividades formativas y sumativas [1]. Gracias al surgimiento del internet, se han dado varios métodos de enseñanza, evaluación y aprendizaje que incluyen el uso de un entorno virtual. Esto ha permitido flexibilizar horarios y espacios de enseñanza para los docentes y espacios de aprendizaje para los estudiantes, donde no se tiene una restricción geográfica o de horario, aumentando el número de estudiantes [2].

para MOOCs. Por ejemplo, el Observatorio Class Central [6] hasta el año 2020 reportó un incremento en más de 110 millones de estudiantes que se han registrado en alguno de los más de 13.500 MOOCs de diferentes áreas de interés, producidos por más de 900 universidades alrededor del mundo y que son responsables de estos contenidos [7]. Por otro lado, el Observatorio UC [8] y el Observatorio MOOC-CEDIA [9] en el reporte anual de 2020 que realizan, destacan a la plataforma Open edX (después de Coursera) como una de las plataformas que más acogida ha tenido durante la pandemia y con una cantidad alta de usuarios en el año 2020 en comparación con años anteriores.

Esto se debe en parte a que la plataforma edX liberó el código fuente de su plataforma denominado Open edX, que cuenta con funcionalidad similar a edX para alojar MOOCs. Sin embargo, Open edX tiene una limitante en relación con las analíticas de datos que se presentan y que se pueden utilizar para extraer conocimiento a partir de las interacciones de los estudiantes con los materiales de la plataforma [10]. Esto ha dado como resultado, que estudiantes y profesores encuentren limitaciones al momento de utilizar los MOOCs más allá de simples “libros digitales” donde la retroalimentación es escasa y el apoyo a los estudiantes es casi nulo [11]. En ese contexto, las Analíticas de Aprendizaje (LA - Learning Analytics en inglés) son una alternativa para mejorar los procesos de enseñanza y aprendizaje por medio del análisis de los datos. Para Trafford y Shirota [12], LA es el campo de estudio que se encarga de medir, recolectar, analizar y reportar datos acerca de los estudiantes y los contextos en los que se desarrolla; y permiten escalar las habilidades de los docentes para dar seguimiento a los estudiantes y retroalimentarlos de forma adecuada.

En un intento por incorporar visualizaciones dotadas de