

**Visualización científica de comportamientos con  
sospecha de fraude en Cursos en Línea Masivos y  
Abiertos (MOOC)**



**Universidad  
del Cauca®**

Trabajo de Grado

**Aldair Luna Buitrón**

**Herman Gabriel Eraso Rosero**

Director: Mag. Daniel Alberto Jaramillo Morillo  
Co-Director: PhD. Mario Fernando Solarte Sarasty

*Departamento de Telemática  
Facultad de Ingeniería Electrónica y Telecomunicaciones  
Universidad del Cauca  
Popayán, Cauca, 2021*

# Visualización científica de comportamientos con sospecha de fraude en Cursos en Línea Masivos y Abiertos

Aldair Luna Buitrón

Herman Gabriel Eraso Rosero

Trabajo de Grado presentado a la Facultad de Ingeniería Electrónica y Telecomunicaciones de la Universidad del Cauca para obtener el título de Ingeniero en Electrónica y Telecomunicaciones

Director: Mag. Daniel Alberto Jaramillo Morillo

Co-Director: PhD. Mario Fernando Solarte Sarasty

*Departamento de Telemática  
Facultad de Ingeniería Electrónica y Telecomunicaciones  
Universidad del Cauca  
Popayán, Cauca, 2021*

# Resumen

Los MOOC (Massive Open Online Course) han causado gran impacto en el mundo académico. Crearon una nueva forma de acceder a la educación de una manera escalable, flexible y asíncrona, y actualmente están siendo implementados y acreditados académicamente dentro de las universidades de educación superior a través de sus variantes SPOC (Small Private Online Course) y MPOC (Massive Private Online Course). Sin embargo, inicialmente los MOOC no fueron pensados para ser certificados y reconocidos, por lo que no ha tenido en cuenta los posibles casos de bajo rendimiento, las altas tasas de abandono y los comportamientos con sospecha de fraude que se presentan en los entornos virtuales. A esto se suma que las diferentes plataformas MOOC no ofrecen las herramientas necesarias para dar apoyo al docente, realizar el seguimiento a los estudiantes y detectar sus comportamientos con sospecha de fraude. El objetivo de este trabajo fue construir un mecanismo de visualización mediante un desarrollo web "SELENEVIS" como una herramienta de apoyo al docente, que le permita visualizar los estudiantes identificados como sospechosos de fraude y explorar el registro de sus actividades. Como resultado se tiene 4 vistas que representan diferentes niveles de abstracción. Una vista de Administrador, una Vista de sospechosos apoyado en el trabajo de [1], una vista de estadísticas de estudiantes y de cursos, y una vista de comparación entre estudiantes a través de una propuesta de esquema de grafos que representan las secuencias de actividades en la plataforma acompañadas con gráficos estadísticos complementarios.

Palabras clave: MOOC, Deshonestidad académica, Visualización, Identificación de sospechosos, Seguimiento, Esquema de grafos.

# Tabla de Contenido

<b>Lista de figuras</b>	<b>viii</b>
<b>1 Introducción</b>	<b>1</b>
1.1 Contexto	1
1.2 Planteamiento del Problema	4
1.3 Objetivos	5
1.3.1 Objetivo General	5
1.3.2 Objetivos Específicos	6
1.4 Metodología	6
1.4.1 WP1. Generación de la base inicial de conocimiento	6
1.4.2 WP2. Modelo de descripción y caracterización de los comportamientos de sospecha de fraude en MOOC	7
1.4.3 WP3. Construir un prototipo de visualización de los comportamientos sospechosos de fraude con base en el modelo de descripción y caracterización propuesto dentro del contexto de los MOOC para la Universidad del Cauca.	7
1.4.4 WP4. Verificar el prototipo con dataset de cursos MOOC de la Universidad del Cauca recopilados en trabajos anteriores.	7

1.4.5 WP5. Publicación . . . . .	8
1.5 Resultados alcanzados . . . . .	8
1.6 Contenido monografía . . . . .	9
<b>2 Conceptos fundamentales</b>	<b>10</b>
2.1 Definición de MOOC . . . . .	10
2.1.1 Masivos (Massive) . . . . .	11
2.1.2 Abiertos (Open) . . . . .	11
2.1.3 En línea (Online) . . . . .	12
2.1.4 Cursos (Courses) . . . . .	12
2.2 Variantes de MOOC . . . . .	12
2.2.1 SPOC . . . . .	13
2.2.2 MPOC . . . . .	13
2.3 Ventajas y desventajas de los MOOC . . . . .	14
2.3.1 Ventajas de los MOOC . . . . .	15
2.3.2 Desventajas de los MOOC . . . . .	16
<b>3 Estado del Arte</b>	<b>18</b>
3.1 Metodología de Investigación . . . . .	18
3.1.1 Definición para la búsqueda . . . . .	19
3.1.1.1 Pregunta de investigación . . . . .	19
3.1.1.2 Alcance de la revisión . . . . .	19
3.1.1.3 Criterios de inclusión y exclusión . . . . .	19

3.1.1.4	Conducta de la búsqueda	20
3.2	Ejecución de la revisión	20
3.3	Trabajos relacionados	22
3.3.1	Reconocimiento académico	22
3.3.2	Deshonestidad académica	23
3.3.2.1	Examen Supervisado	23
3.3.2.2	Análisis de aprendizaje	24
3.3.3	Seguimiento al estudiante	26
3.3.3.1	Recolección de datos	26
3.3.3.2	Visualización	28
<b>4</b>	<b>Mecanismo de visualización de datos</b>	<b>31</b>
4.1	Recolección de datos	31
4.1.1	Plataforma de aprendizaje	32
4.1.1.1	Open Edx	32
4.1.1.2	Selene Unicauca	32
4.1.2	Características Básicas Open Edx	34
4.1.2.1	Sistema de Gestión de Aprendizaje	35
4.1.2.2	Sistema de Gestión de Contenidos	35
4.1.2.3	Tracking.log	36
4.1.3	Extracción de datos desde Selene Unicauca	36
4.2	Caracterización de datos	37

---

4.2.1	Caracterización de los datos	38
4.2.2	Construcción de grafos	38
4.2.2.1	Presentación de Grafos	42
4.3	Visualización	43
4.3.1	Arquitectura general	43
4.3.1.1	Módulo Recolector datos	43
4.3.1.2	Módulo Detector sospechosos	44
4.3.1.3	Módulo Estadísticas	44
4.3.1.4	Módulo Grafos	45
4.3.1.5	Módulo Admin	45
4.3.1.6	Módulo Vistas	45
4.4	Descripción arquitectura SELENEVIS	46
4.4.1	Vista de escenarios	46
4.4.2	Vista lógica	48
4.4.3	Vista de procesos	50
4.4.4	Vista de Desarrollo	54
4.4.5	Vista física	54
4.4.6	Implementación	55
4.4.6.1	Repositorio	55
4.4.6.2	Servidor	56
4.4.6.3	Repositorio	56

4.4.6.4	GitLab Runner	57
4.4.6.5	Servidor HTTP	58
4.4.6.6	Dominio SELENEVIS	59
4.5	Caso de estudio	59
4.5.1	Descripción del curso	60
<b>5</b>	<b>Resultados</b>	<b>62</b>
5.1	SeleneVis	62
5.1.1	Vista “Login”	62
5.1.2	Vista “Sospechosos”	63
5.1.3	Vista “Estadísticas”	65
5.1.4	Vista “Comportamientos”	67
5.1.5	Vista “Admin”	69
5.2	Resultados sobre el Caso de estudio	69
5.2.1	Identificación de sospechosos de fraude en SELENEVIS	71
5.2.1.1	Estadísticas generales sospechosos	79
5.2.1.2	Gráficas línea de tiempo	81
5.2.1.3	Gráfica diferencia de tiempos en envíos	85
5.2.1.4	Esquema de Grafos grupo 3, 4 y 7	86
5.2.1.5	Conclusiones	94
<b>6</b>	<b>Conclusiones</b>	<b>96</b>



---

<b>7 Trabajos Futuros</b>	<b>100</b>
<b>Bibliografía</b>	<b>100</b>
<b>Anexos</b>	<b>113</b>
<b>A Manual de Usuario SeleneVis</b>	<b>113</b>
A.1 Vista login . . . . .	113
A.2 Vista estadísticas . . . . .	114
A.3 Vista sospechosos . . . . .	115
A.4 Vista comportamientos . . . . .	116
A.5 Vista administración . . . . .	117
<b>B Manual de instalación Selenevis</b>	<b>119</b>
B.1 Dependencias . . . . .	119
B.2 Clonar repositorio . . . . .	119
B.3 Instalar dependencias . . . . .	120
B.4 Generar archivos producción . . . . .	120
B.5 Configurar base de datos . . . . .	120
B.6 Lanzar aplicación . . . . .	121
<b>C Artículo</b>	<b>122</b>

# Lista de Figuras

3.1 Resultados totales de búsqueda	20
3.2 Resultados de búsqueda con relevancia	21
3.3 Número de publicaciones por año	21
4.1 Logo Open Edx. Tomada de [2]	32
4.2 Selene Unicauca Inicio	33
4.3 Arquitectura para recolección de datos	36
4.4 Estructura de Registro JSON	37
4.5 Caracterización de eventos	39
4.6 Ejemplo control de aristas y nodos	40
4.7 Creación de grafos	40
4.8 Estructura de grafos	41
4.9 Representación visual de grafos	42
4.10 Arquitectura General Mecanismo SeleneVis. Adaptada de [3]	44
4.11 Modelo de 4+1 vistas del Proceso Unificado de Desarrollo (Rational Unified Process, RUP).	46

---

4.12 Diagrama de casos de uso	47
4.13 Diagrama de clases general	48
4.14 Diagrama de clases SELENEVIS	49
4.15 Diagrama de Proceso-Recolección De Datos	50
4.16 Diagrama de Proceso-Detección de Sospechosos	52
4.17 Diagrama de Proceso-Módulo de Visualización	53
4.18 Diagrama de Desarrollo Mecanismo de Visualización	54
4.19 Vista física	55
4.20 Repositorio en GitLab	57
4.21 GitLab Runner	58
4.22 Login SeleneVis	59
5.1 Vista de login	63
5.2 Vista sospechosos. Resumen general	64
5.3 Similitud estudiantes	65
5.4 Vista estadísticas	66
5.5 Estadísticas por estudiante	67
5.6 Vista comportamientos	68
5.7 Vista de Administrador	69
5.8 Curso LeanStartUp	70
5.9 Resumen general	72
5.10 Unidad temática I	73

---

5.11 Sospechosos con porcentaje de similitud obtenidos de la aplicación . . .	74
5.12 a. Porcentaje de similitud b. Diferencia de tiempo (segundos) . . . . .	76
5.13 Caracterización sospechosos por nota de examen y similitud . . . . .	77
5.14 Comparativo general de estadísticas Grupos 3, 4 y 7 . . . . .	79
5.15 Comparativo estadísticas grupo 3, 4 y 7 . . . . .	80
5.16 Interacción en línea de tiempo grupo 3 . . . . .	82
5.17 Interacción en línea de Tiempo grupo 4 . . . . .	83
5.18 Interacción en línea de tiempo grupo 7 . . . . .	84
5.19 Diferencia de tiempos de envío grupos 3, 4 y 7 . . . . .	85
5.20 Grupo 3 examen 3 S5, S6 . . . . .	88
5.21 Grupo 3 examen 3 S7 . . . . .	88
5.22 Grupo 3 examen 4 S5, S6 . . . . .	89
5.23 Grupo 3 examen 4 S7 . . . . .	90
5.24 Grupo 4 examen 3 . . . . .	91
5.25 Grupo 4 examen 5 . . . . .	92
5.26 Grupo 7. Examen 3 . . . . .	93
5.27 Grupo 7 Examen 5 . . . . .	94
A.1 Vista de login . . . . .	113
A.2 Vista estadísticas . . . . .	114
A.3 Vista sospechosos . . . . .	115
A.4 Vista comportamientos . . . . .	116

---

A.5 Vista administración	117
B.1 Elaboración propia. Crear usuario administrador	121

# Capítulo 1

## Introducción

### 1.1 Contexto

Los cursos en línea masivos y abiertos (MOOC, Massive Open Online Course) han causado gran impacto en la educación en línea incluso entre las más prestigiosas universidades del mundo (Stanford, Harvard, MIT, entre otros), de tal forma que en noviembre de 2012 el periódico New York Times publicó un artículo llamado “El año de los MOOC” (The Year Of The MOOC) haciendo referencia a la gran atención generada por este término sobre la comunidad universitaria y los medios de comunicación. Sin embargo, los cursos MOOC fueron referenciados por George Siemens y Stephen Downes en la Universidad de Manitoba (Canadá) en agosto de 2008, a través del curso que llevó por nombre “Conectivismo y conocimiento conectivo” (Connectivism and Connective Knowledge), en el que se inscribieron 2.300 participantes de todo el mundo, lo que impulsó a los autores a la creación de nuevos cursos [4, 5, 6].

El camino hacia el éxito de los MOOC se materializó con dos casos rotundos, el curso de “Introducción a la inteligencia artificial” (Introduction to Artificial Intelligence), organizado en otoño de 2011 por Sebastian Thrun, profesor de la Universidad de Stanford y Peter Norvig, curso que contó con 160.000 participantes de todas partes del mundo [5, 7]. Poco después tuvo lugar el curso de “circuitos y electrónica”

(Circuits and Electronics) organizado por el profesor Anant Agarwal del Instituto de Tecnología de Massachusetts (MIT) en su plataforma MITx, con más de 120.000 estudiantes inscritos de todo el mundo en la primavera de 2012. En vista del gran éxito en los MOOC, el profesor Sebastian Thrun abandonó su puesto de docente en la Universidad de Stanford para crear una plataforma de soporte para el despliegue de nuevos MOOC a la que llamó Udacity [8]. Así mismo los profesores Andrew Ng y Daphne Koller de la universidad de Stanford, fundaron la plataforma Coursera y comenzaron a ofrecer cursos a partir de abril de 2012 [9].

Tras estos ejemplos, los MOOC tuvieron gran trascendencia, de tal forma que dos universidades líderes en calidad de educación, como son Harvard y MIT se unieron para invertir y conjuntamente formar un proyecto sin ánimo de lucro llamado Edx [10], cuyo propósito era difundir los recursos educativos y romper los esquemas de la educación tradicional.

Es así como en el 2012 se llevó los MOOC al público en el que la comunidad empresarial y universitaria han realizado grandes esfuerzos, dando lugar a la aparición de muchas plataformas para el despliegue y soporte de MOOC, entre ellas: Coursera, Udacity, edX, MiríadaX, Udemy, Khan Academy, Codecademy, Lynda, SkilledUp, Academic Earth, Saylor.org, Canvas Network, entre otras [11] [12].

En la actualidad, los MOOC muestran números asombrosos. Al buscar MOOC en internet se despliegan millones de resultados. El crecimiento del número de participantes se ha hecho evidente, lo que a principios se medía en cientos de estudiantes, para el 2017 se cuantificó en millones de estudiantes inscritos en cursos MOOC en cientos de universidades del mundo. Los números no han parado de crecer y para 2018 se tuvo más de 111 millones de participantes registrados en las plataformas MOOC y cerca de 11.000 cursos en más de 900 universidades. Según los datos publicados por Class Central, un motor de búsqueda especializado en educación en línea [13], [14].

No todos los escenarios son tan favorables para los MOOC, si bien su carácter masivo y abierto son pilares fundamentales en el éxito adquirido, también ha generado dificultad a la hora de un aprendizaje personalizado. Existe una enseñanza pedagógica rígida y la interacción entre alumno docente se dificulta a partir de las herramientas

existentes dentro de las plataformas. Esto acarrea un problema de deserción muy elevado, solo del 6% al 7.5% de los inscritos termina el curso. Muchas razones se han estudiado, entre ellas; el hecho de que en la mayoría de los casos la inscripción al curso sea gratuita, que los contenidos sean poco personalizados, que las personas busquen un aprendizaje más específico, la falta de control y seguimiento del comportamiento de los estudiantes [15].

Sin embargo desde otra perspectiva se han realizado estudios en los que se ha podido identificar que la obtención de certificaciones e insignias es de mucha importancia para los participantes a la hora de completar los cursos. Por esta razón los diferentes entes de educación han hecho grandes esfuerzos en incluir o validar estos reconocimientos tanto dentro de sus créditos educativos como en diferentes contextos laborales a través de certificados, créditos educativos e insignias digitales. Un ejemplo de ellos serían: Los Nanodegrees de Udacity, los XSeries y los MicroMasters de edX, los programas especializados y certificados MasterTrack de Coursera y los certificados de graduación y diplomas de FutureLearn, entre otros [16, 7, 17].

Si bien el reconocimiento académico influye positivamente en los MOOC, trae consigo una nueva preocupación dado que inicialmente los MOOC no se consideraron para ser reconocidos o acreditados, por lo tanto, no se implementó herramientas de seguimiento a los comportamiento de los estudiantes y por lo tanto no existe una relación adecuada entre el profesor y el estudiante y se generan preocupaciones en cuanto a la honestidad académica de los estudiantes; saber si los estudiantes son los mismos quienes completan los cursos o incluso saber si los estudiantes no comenten fraude. De cualquier manera, poco a poco los MOOC se han ido incorporando al ambiente universitario a través de los SPOC (Small Private Online Course) y los denominados MPOC (Massive Private Online Course), que consideran grupos más pequeños y selectivos dado su característica de privacidad y por tanto han permitido implementar algunos métodos de supervisión dependiendo del tipo de reconocimiento académico adquirido tras la finalización del curso, ya sea reconocimiento de participación, superación o de convalidación (reconocimiento académico) [18].



## 1.2 Planteamiento del Problema

Los MOOC poseen ventajas en cuanto a flexibilidad y escalabilidad. Flexibilidad en cuanto vencen las barreras espacio temporales y admiten el acceso al material en línea en cualquier momento y lugar; lo que permite la autonomía de los estudiantes con las actividades del curso y abre una puerta de acceso a los recursos de prestigiosas universidades del mundo como la Universidad de Oxford, Cambridge, Stanford, entre otros. Cabe resaltar que en su gran mayoría el contenido impartido proviene de docentes de talla mundial. La escalabilidad por otra parte hace referencia a la posibilidad de establecer nuevas relaciones, contar con una excelente calidad de la información, propiciar el trabajo colaborativo; y por su gran cantidad de participantes, generar nuevos aportes que contribuyen al proceso formativo de todos [13], [19].

Sin embargo, en medio de las grandes ventajas que poseen los MOOC, existen desafíos por superar asociados a la escalabilidad de las plataformas, en gran medida relacionados a la falta de interacción directa entre profesor y estudiante. Al presentarse una gran cantidad de participantes se dificultan tareas como la evaluación, el seguimiento y la retroalimentación, provocando pérdida de interés de los estudiantes y altas tasas de deserción [15], [20].

Además de la poca interacción entre profesores y estudiantes, también se derivan fallos relacionados a la deshonestidad académica que se presenta como consecuencia del escaso control en el desarrollo de exámenes y otras actividades de evaluación, siendo este uno de los principales inconvenientes cuando los cursos tienen reconocimiento académico y su respectiva certificación [20], [21], [15].

En el contexto de los MOOC, la inclusión de una certificación representa un factor motivacional para validar la adquisición de nuevos conocimientos y competencias en el sector laboral y de la educación. Por esta razón el reconocimiento académico de los MOOC está siendo impulsado por las principales plataformas (Coursera, EdX, Udacity, Miriadax, RedunX, UniMOOC, entre otros), algunas de las universidades quienes han creado MOOC se han propuesto validarlos y reconocerlos dentro de sus programas profesionales. Sin embargo, no han tenido en cuenta los comportamientos

deshonestos que pueden aparecer en los estudiantes, por lo que varios autores han manifestado una constante incertidumbre de fraude al momento de validarlos en programas universitarios [18].

Algunos trabajos han recopilado y analizado los datos generados de la interacción de los estudiantes con foros, cadenas de clics en vídeos, dispositivos con los que accede, cantidad de módulos desarrollados, número de accesos a la plataforma o el tiempo durante el desarrollo del examen; con el fin de entender mejor este modelo. Sin embargo, no se centran en la manera en cómo estos datos son presentados al docente o en cómo un docente puede estar al tanto de los comportamientos deshonestos que puede tener un estudiante en un curso. Aún existe la necesidad de brindar una herramienta de apoyo al docente con el fin de realizar un seguimiento que permita visualizar el comportamiento deshonesto de los estudiantes en cursos reconocidos y acreditados académicamente [18], [22], [23], [24], [25], [26], [27], [28].

Por tanto, surge la siguiente pregunta de investigación: ¿Cómo visualizar el comportamiento con sospecha de fraude del estudiante en un MOOC?

Tomando como referencia el trabajo realizado por el estudiante de doctorado Daniel Jaramillo Murillo [1], se plantea, crear un modelo gráfico que permita visualizar el comportamiento de los estudiantes con sospecha de fraude allí encontrados.

## 1.3 Objetivos

A continuación se describen los objetivos planteados para el desarrollo del proyecto de investigación:

### 1.3.1 Objetivo General

Proponer un modelo para la visualización de comportamientos sospechosos de fraude en cursos en línea abiertos y masivos MOOC.

### 1.3.2 Objetivos Específicos

- Caracterizar los datos <sup>1</sup> del comportamiento de los estudiantes con sospecha de fraude en MOOC.
- Desarrollar un mecanismo (desarrollo software) que permita la visualización de los datos de comportamiento <sup>2</sup> de los estudiantes en MOOC.
- Verificar el modelo mediante la puesta en marcha de un caso de estudio en el contexto de la Universidad del Cauca.

## 1.4 Metodología

Para el desarrollo de las actividades del cronograma y para dar cumplimiento a los objetivos planteados, se usó como referencia la descomposición jerárquica WBS (Work Breakdown Structure) sugerida en la metodología PMBOK (Project Management Base of Knowledge) por el PMI (Project Management Institute), específicamente en el área de gestión del alcance (Scope Management) <sup>[29]</sup>.

A continuación se muestran las actividades para cada uno de los paquetes de trabajo propuestos.

### 1.4.1 WP1. Generación de la base inicial de conocimiento

- Definición del objetivo de la vigilancia (planeación)
- Construcción del corpus (búsqueda y captación)
- Análisis del corpus (análisis)

---

<sup>1</sup>Datos recopilados de la base de datos construida en el trabajo de grado "Seguimiento automático de las actividades de los estudiantes en un MOOC", descrita en <sup>[18]</sup>, la cual recolecta información de la interacción de los estudiantes en un curso MPOC de la plataforma Selene de la Universidad del Cauca.

<sup>2</sup>Aquellos comportamientos con sospecha de fraude identificados a partir de la caracterización de los datos de interacción de los estudiantes con la plataforma en el desarrollo de un curso MOOC

### **1.4.2 WP2. Modelo de descripción y caracterización de los comportamientos de sospecha de fraude en MOOC**

- Búsqueda de modelos existentes para la descripción de los comportamientos con sospecha de fraude en cursos en línea.
- Análisis de los modelos y su compatibilidad en el ámbito de los MOOC.
- Diseño del modelo que caracterice los comportamientos con sospecha de fraude en MOOC.

### **1.4.3 WP3. Construir un prototipo de visualización de los comportamientos sospechosos de fraude con base en el modelo de descripción y caracterización propuesto dentro del contexto de los MOOC para la Universidad del Cauca.**

- Análisis de las técnicas que permitan la construcción de un prototipo para la visualización de las variables propuestas en el modelo de caracterización.
- Elección de la técnica apropiada.
- Implementación de la técnica sobre un dataset previamente obtenido de la plataforma Open edX.

### **1.4.4 WP4. Verificar el prototipo con dataset de cursos MOOC de la Universidad del Cauca recopilados en trabajos anteriores.**

- Diseño y desarrollo del caso de estudio.
- Estructuración de los datos del dataset.

- Selección e implementación de los datos que permitan visualizar el comportamiento de los estudiantes en un MOOC.

#### 1.4.5 WP5. Publicación

- Generación de artículos sobre los avances de la investigación
- Generación del documento final (monografía)

### 1.5 Resultados alcanzados

- Síntesis del estado del arte donde se muestra diferentes estudios aplicados al ámbito de los MOOC, con el fin de descubrir el comportamiento del estudiante dentro de un curso.
- Documentación de la identificación de las variables existentes en el archivo tracking-log de la plataforma Selene Unicauca instancia de Open-Edx.
- Diseño de una arquitectura para la implementación de una aplicación web que permita procesar, guardar y visualizar los datos que permiten determinar los comportamientos de los estudiantes de un MOOC.
- Implementación de una aplicación web, la cual permite ver los datos estadísticos de los cursos en la plataforma, visualizar los comportamientos de los estudiantes, realizar comparaciones entre estudiantes y visualizar a los estudiantes con sospecha de fraude dentro de un MOOC.
- Realización de pruebas de la aplicación web con base en la información obtenida de un curso anterior "LeanStartUp 2020-I".
- Implementación de un caso de estudio con base en la aplicación web desarrollada.

## 1.6 Contenido monografía

Capítulo 1 - INTRODUCCIÓN.

Definición del problema y la estructura general del desarrollo del proyecto de investigación.

Capítulo 2. CONCEPTOS FUNDAMENTALES.

Presenta un marco teórico en el que se muestra la terminología y conceptualización empleada en el desarrollo del proyecto.

Capítulo 3. ESTADO DEL ARTE.

En el cual se presenta un resumen de los trabajos relacionados existentes y experiencias previas de otras investigaciones acerca de modelos de visualización que permitan la identificación de comportamientos con sospecha de fraude en entornos MOOC.

Capítulo 4. MECANISMO DE VISUALIZACIÓN DE DATOS

Se ilustra el proceso de diseño e implementación de un prototipo basado en una aproximación arquitectónica que soporte la visualización de los comportamientos con sospecha de fraude en base al conjunto de identificadores encontrados.

Capítulo 5. RESULTADOS

En donde se exponen los resultados alcanzados con el uso de la aplicación “SELENEVIS”, además de analizar el curso piloto y las diferentes gráficas obtenidas.

Capítulo 6. CONCLUSIONES.

Se analizan los resultados del trabajo realizado, se detallan las principales contribuciones obtenidas durante el ciclo del proyecto y se expone un conjunto de recomendaciones importantes para el desarrollo de trabajos futuros.

# Capítulo 2

## Conceptos fundamentales

En este capítulo se realiza una descripción de los conceptos claves para la realización del proyecto.

### 2.1 Definición de MOOC

Una de las innovaciones tecnológicas más importantes de los últimos 200 años dentro del ámbito educativo, son los MOOC, ya que ofrecen un aprendizaje flexible reuniendo a docentes y aprendices de todo el mundo con ideas afines. Este tipo de cursos se basa en un enfoque conectivista. lo cual permite la distribución de toda la información a través de la red de participantes [30].

Las siglas MOOC "Massive Open Online Course" hacen referencia a una de las tendencias más prometedoras, ya que no solo representan contenido educativo en vídeos o exámenes al igual que los cursos tradicionales, sino que además, tienen un componente extra que son los foros, los cuales son un espacio en el que los estudiantes pueden poner en discusión diferentes temas de importancia [31].

Los MOOC, son cursos en línea masivos y abiertos, alojados en una plataforma virtual, que al igual que los cursos tradicionales tiene un inicio y un fin, un docente encargado y actividades de evaluación. Sin embargo estos cursos no tienen un

número limitado en cuanto a participantes y los horarios de estudio son flexibles pues el estudiante es quien decide en que momento accede a la plataforma para tomar las clases y realizar las actividades de evaluación [32].

Las características de estos cursos son desglosadas a continuación:

### 2.1.1 Masivos (Massive)

Los MOOC, adoptan la característica de ser masivos dado que están dirigidos a una cantidad mayor de estudiantes comparados con los cursos tradicionales, además, gracias a las plataformas donde son implementados pueden llegar a más personas alrededor del mundo, es decir, no se tiene un límite de participantes y cualquier persona con conexión a internet puede acceder a ellos [33].

Por otro lado, dado que los cursos deben cumplir con esta característica es importante resaltar que las plataformas y el software involucrado debe ser escalable para que no se presenten inconvenientes en la accesibilidad de la información, ya que no existe límite en las inscripciones de los estudiantes. Además, para garantizar la masividad del contenido, los cursos debe ser alcanzables a nivel global, es decir, se debe permitir el acceso a estudiantes de todo el mundo sin importar el idioma u horario [32].

### 2.1.2 Abiertos (Open)

Otra característica importante de los MOOC, es su capacidad de ser abiertos, ya que ofrecen una gran cantidad de material educativo en forma de vídeos, foros, PDF, multimedia, test, entre otros. Dicho material está disponible a lo largo del curso abiertamente para que el conocimiento esté a disposición de los estudiantes y docentes y pueda ser aprovechado eficientemente [30], [34].

Es importante resaltar, que la característica de ser abiertos hace posible que diversas personas puedan participar de los cursos, a partir de una inscripción o registro previo sin ningún tipo de condicionamientos para aprender o acceder a los contenidos. Al ser “Abiertos” refuerza el sentido que tienen los MOOC, ya que su filosofía se basa



en la circulación del conocimiento generalmente de manera gratuita para que más personas puedan acceder abiertamente [35].

### 2.1.3 En línea (Online)

Los MOOC, desde su filosofía son cursos que están disponibles en plataformas que se pueden soportar en línea, es decir están dispuestas para ser atendidas y accedidas desde cualquier lugar geográfico a partir de la red de Internet [36].

La característica “Online” significa que tanto el aula, el docente y el contenido están en un ambiente virtual; por lo tanto, la asistencia a un aula física y la interacción cara a cara con el docente queda de lado. Al presentar esta característica, el estudiante puede flexibilizar sus horarios de estudio, logrando acceder al contenido en cualquier momento pues está disponible las 24 horas del día [37].

### 2.1.4 Cursos (Courses)

A pesar de que se desarrolla en un ambiente virtual, es más que necesario que un docente o grupo de docentes planifiquen y definan la estructura del curso de una manera transparente y entendible para todos los estudiantes. Además estos cursos deben contar con un cronograma de actividades y exámenes. También es posible que estos cursos sean complementados con sesiones presenciales; si es viable dentro del contexto en el que se desarrollan. De tal manera que se disminuya la brecha existente en cuanto a la acreditación real del conocimiento y las relaciones entre estudiantes y docentes [38], [39].

## 2.2 Variantes de MOOC

A medida que la implementación de los cursos MOOC se ha ido incrementando en organizaciones y universidades, algunas variantes han surgido con el objetivo de suplir algunos requerimientos que los MOOC originales no pueden hacer. La

implementación de estas variantes de MOOC generan innovación en la forma de enseñar dentro de las universidades [1].

En esta sección, se describe las principales variantes existentes en la actualidad.

### 2.2.1 SPOC

Los SPOC son una variante de los MOOC, que se crearon bajo la necesidad de las diferentes universidades de ajustar estos cursos a la población estudiantil y bajo las capacidades de escalabilidad de cada institución. Es decir, aprovechando las bondades de los MOOC, pero en grupos más pequeños, permitiendo no solo personalizar el aprendizaje sino un mayor acercamiento en la interacción alumno-docente [40]. Visto desde otra perspectiva, se busca un balance entre la dinámica de interacción del estudiante con los recursos de calidad de un curso MOOC, y la dinámica de la enseñanza tradicional de un aula de clase. Un SPOC en el objetivo de mejorar la calidad de enseñanza y sostenibilidad, se diferencia de un Curso MOOC tanto en tamaño y accesibilidad [41].

El “Small” de su acrónimo representa que un SPOC no es de carácter masivo, sino que su número de participantes alcanza unas pocas decenas o hasta unos cientos de estudiantes. Por otra parte, “Private” hace referencia a que el acceso no es completamente abierto a todo el mundo, sino a un grupo selecto de personas, como puede ser un curso dispuesto por una universidad solo para sus estudiantes, que puede ser gratuito o de pago. A partir de estas características, un SPOC puede soportarse sobre instancias de plataformas MOOC, como es el caso de Open Edx o Moodle, sobre los que las Instituciones despliegan sus cursos personalizados [42], [43], [44].

### 2.2.2 MPOC

Los MPOC (Massive Private Online Course), son una variante de los MOOC tradicionales que se distinguen por estar restringidos a través de una característica de privacidad, es decir los MPOC no están completamente abiertos a todo público ni son necesariamente

gratuitos, por lo que el número de estudiantes inscritos no es tan grande como en los MOOC, pero si suelen tomar números mayores de participantes en contraste con la cantidad de aprendices inscritos en los cursos en línea convencionales o los SPOC [45].

El término MPOC aparece en 2014, como fruto del esfuerzo de la Universidad de Pekín en la implementación y desarrollo de un curso de formación docente para un grupo selecto de maestros (proyecto K-12), en los que se buscaba dar capacitación de la inclusión de las nuevas Tecnologías de la Información y Comunicación (TIC) para el mejoramiento de la calidad de la educación básica en las diferentes provincias de China. El curso se diseñó según el modelo de un SPOC, con clases en línea para unas pocas decenas de participantes a las cuales se asigna un facilitador en línea que orienta el aprendizaje. Aunque se presentaron algunos prejuicios en China sobre los cursos en línea y el bajo costo de matrícula, el número real de participantes alcanzó un número de cientos de inscritos. Aproximadamente, entre 2007 y 2013, cerca de 500,000 maestros de las diferentes provincias se inscribieron, dando lugar a 5,000 salas en línea, para lo cual se añadieron cerca de 2,000 facilitadores. Fue con esto que se descubrió una nueva práctica de capacitación MPOC para maestros en China [46].

## 2.3 Ventajas y desventajas de los MOOC

La incorporación de los cursos MOOC en el modelo de aprendizaje se ha incrementado en los últimos años, y la entrega de contenido de los cursos se ha relacionado con múltiples ventajas, ya que la virtualidad junto con la inclusión del estudiante en los procesos promueve un aprendizaje óptimo. Sin embargo, así como existen múltiples ventajas también se presentan algunas desventajas identificadas a partir de los estudios realizados hasta el momento para este tipo de cursos [47]. Las ventajas y desventajas son descritas a continuación:

### 2.3.1 Ventajas de los MOOC

Los MOOC ofrecen diversas ventajas a los estudiantes, docentes e instituciones de acuerdo a los siguientes aspectos tomados a partir de [48], [49], [50]:

- Acceso a la educación: los estudiantes que participan de los cursos MOOC son variados y abundantes, lo que indica que se incrementa la accesibilidad a la educación. Por un lado, se cuenta con la característica de que estos cursos son gratuitos y por el otro que son masivos, virtuales y abiertos; lo que facilita el acceso de personas que no pueden incurrir en una carga financiera elevada y/o que se encuentran en áreas geográficas lejanas de centros educativos. Por lo tanto, la audiencia de estos cursos es global.
- Flexibilidad: Es importante resaltar que los estudiantes obtienen flexibilidad en su educación cuando se involucran en un curso MOOC, ya que pueden disponer de su tiempo para acceder y completar el contenido del curso, evitando las preocupaciones de cumplir con un horario establecido. El estudiante puede realizar diferentes actividades, como cumplir con sus obligaciones de trabajo, familia u otras.
- Proactividad: los MOOC, permiten que el estudiante vaya a su ritmo dado que tiene un acceso libre al contenido, al mismo tiempo facilita la interacción con otros estudiantes para generar un intercambio de conocimientos. Además, un estudiante puede participar de foros u otras vías de comunicación para resolver sus dudas, participar o exponer su punto de vista de acuerdo al contenido del curso.
- Aprendizaje permanente: el crecimiento profesional de un estudiante se forja a partir de los conocimientos que vaya adquiriendo progresivamente. El estudiante puede acceder a una cantidad considerable de material de diversos temas dado que se suben cursos nuevos continuamente y los cursos existentes son actualizados, por lo cual, este modelo es ideal para estudiantes motivados en crear su propio camino educativo.

- Reducción de costos: debido a la significativa interacción social que presentan los MOOC, el apoyo y la colaboración entre los participantes del curso permiten una reducción de costos, debido a que las distintas subtarefas del proceso de enseñanza que normalmente son costeadas por el proveedor de servicio educativo, las desarrollan los mismos participantes.
- Certificación: en caso que el estudiante requiera una certificación del curso aprobado, puede solicitarlo. Lo único que se debe considerar es que el tema de la certificación no es gratuito, por lo cual, tendrá que incurrir con los gastos requeridos.
- Calidad: muchos de los estudiantes que participan de un MOOC, identifican que la calidad del curso es mucho mejor que la de un curso tradicional, debido a que los docentes se ven desafiados a crear contenido de una manera creativa y comprensible.

### 2.3.2 Desventajas de los MOOC

A pesar que los MOOC tienen un gran potencial dentro de la educación, aún presentan grandes desventajas expuestas a continuación [51], [49], [47]:

- Falta de seguimiento al estudiante: el docente no puede hacer un seguimiento detallado de las actividades de los estudiantes en la plataforma.
- Carencia de supervisión de exámenes: debido a que es un curso en línea existen múltiples métodos para responder a la evaluaciones de manera deshonestas, por ejemplo no es posible saber con exactitud si el estudiante que está registrado en el curso es quien realiza el examen o con que grado de honestidad realiza la actividad. Además, también se pueden presentar riesgos de seguridad en Internet.
- Alta tasa de deserción: la mayoría de participantes de MOOC ven estos cursos como un complemento de enseñanza y adquisición de nuevas habilidades, no como el medio principal de aprendizaje, lo que hace que la mayoría de estudiantes abandonen o pierdan el interés y no finalicen el curso.

- **Analfabetismo informático:** es importante que los estudiantes que desean acceder a los cursos MOOC posean como mínimo, un nivel básico de habilidades informáticas, para que se puedan desenvolver sin ningún inconveniente durante el desarrollo del curso. Muchas personas no cuentan con estas habilidades, por lo cual, no se sienten motivadas a iniciar un proceso de aprendizaje mediante MOOC.
- **Tecnología educativa:** es importante que los docentes encargados de desarrollar el contenido del curso cuenten con un equipo tecnológico adecuado. Con frecuencia este es un obstáculo para ellos, pues no reciben incentivos o suficiente apoyo y herramientas, por parte de las instituciones para desarrollar este tipo de cursos.
- **Efectividad de la oferta:** se refiere a que la oferta de cursos MOOC es limitada en campos específicos de la educación cuando se requieren desarrollos prácticos o de laboratorio, dado que aún no se pueden desarrollar dentro de un entorno virtual.
- **Falta de sostenibilidad financiera:** actualmente el negocio financiero se basa en la compra de certificados, tutorías o acreditaciones.

# Capítulo 3

## Estado del Arte

Este capítulo presenta los resultados obtenidos de la aplicación de una Revisión Sistemática para la generación y síntesis del estado actual de conocimiento, donde se muestran los principales trabajos relacionados y empleados en el desarrollo del proyecto. El estado del arte ha permitido encontrar brechas existentes en cuanto a seguimiento en MOOC y además, ha permitido la identificación de las estrategias evaluativas presentadas en este documento.

### 3.1 Metodología de Investigación

Para la generación de la base del conocimiento y el establecimiento de aspectos importantes para satisfacer los objetivos del proyecto, se tomó como base la Vigilancia Tecnológica que es un proceso que permite captar, analizar y difundir información que sirva como base para tomar decisiones importantes con el objetivo de minimizar las amenazas que se puedan presentar durante el desarrollo de un proyecto. La revisión sistemática consta de 3 fases: definición para la búsqueda, ejecución de la búsqueda y discusión de los resultados [52]. A continuación, se describen las tres primeras fases empleadas para la metodología de investigación:

### 3.1.1 Definición para la búsqueda

#### 3.1.1.1 Pregunta de investigación

La definición para la búsqueda será la misma de la pregunta de investigación. ¿Cómo visualizar el comportamiento con sospecha de fraude del estudiante en un MOOC?

#### 3.1.1.2 Alcance de la revisión

Este trabajo se basa en la búsqueda de trabajos científicos en las bibliotecas digitales. Las bases de datos usadas en este trabajo son: Science Direct, Scopus, Ebsco, IEEE Xplore y Springer Link.

Las cadenas de búsqueda en las bases de datos digitales fueron:

( (“data”) AND (“visualization”) AND (“mooc”)) ( (“behaviour”) AND (“mooc”))  
( (“dishonesty”) AND (“mooc”))

#### 3.1.1.3 Criterios de inclusión y exclusión

Para filtrar los resultados se aplicaron los siguientes filtros:

- Son elegibles todos los artículos los cuales tengan relación con la visualización de datos MOOC.
- Son elegibles todos los trabajos los cuales tengan relación con la deshonestidad académica en MOOC.
- Se incluyeron artículos en idiomas inglés y español.
- Se incluyeron artículos publicados entre el año 2010 y 2020.



### 3.1.1.4 Conducta de la búsqueda

Para la selección de los artículos se lleva a cabo los siguientes filtros de revisión.

- **Primer filtro:**
  - **Título:** Se revisan los títulos de las publicaciones arrojadas en las bases de datos.
  - **Resumen o Abstract:** después de revisar los títulos se hace la lectura y análisis del resumen del artículo
- **Segundo filtro:** finalmente se hace una lectura y análisis de los trabajos seleccionados.

## 3.2 Ejecución de la revisión

### Selección de estudios primarios

Se realiza una tabla inicial con el resultado total de las búsquedas y posteriormente a partir de dicha tabla se realiza un análisis para obtener los artículos más relevantes e influyentes para el desarrollo del proyecto.

La figura [3.1](#) muestra los resultados relacionados con el proceso de búsqueda de los documentos relacionados con las palabras clave y fecha de consulta a 28 de marzo de 2021.

RESULTADO TOTAL DE BÚSQUEDA					
Búsqueda	Science Direct	Scopus	Ebsco	IEEE Explore	Springer Link
(("data") AND ("visualization") AND ("mooc"))	409	59	195	57	174
(("behaviour") AND ("mooc"))	1039	563	21	173	514
(("dishonesty") AND ("mooc"))	9	5	2	1	16

Figura 3.1 Resultados totales de búsqueda.

La figura 3.2 muestra el número de artículos que se seleccionaron de acuerdo a cada biblioteca digital y a los términos de búsqueda.

RESULTADO TOTAL DE BÚSQUEDA					
Búsqueda	Science Direct	Scopus	Ebsco	IEEE Explore	Springer Link
( "data" ) AND ( "visualization" ) AND ( "mooc" )	50	25	0	10	0
( "behaviour" ) AND ( "mooc" )	10	5	1	4	3
( "dishonesty" ) AND ( "mooc" )	1	2	0	1	1

Figura 3.2 Resultados de búsqueda con relevancia.

En la figura 3.3 se puede observar el crecimiento en el número de publicaciones año tras año referente a los MOOC.



Figura 3.3 Número de publicaciones por año.

### 3.3 Trabajos relacionados

En las subsecciones [3.3.1](#), [3.3.2](#) y [3.3.3](#) se describen los trabajos encontrados, los cuales fueron divididos en categorías como: reconocimiento académico, deshonestidad académica y seguimiento al estudiante. Se da una breve descripción del trabajo realizado y cuales fueron sus resultados.

#### 3.3.1 Reconocimiento académico

Varios estudios se han esforzado en dar validación a las credenciales digitales y certificados en el contexto de los MOOC, además de mostrar su relación con las tasas de deserción por parte de los estudiantes en los cursos. Los estudios en este ámbito han arrojado algunas percepciones favorables. El caso de [53](#) es una experiencia de un MOOC sobre los derechos humanos, donde se pretendía explorar la viabilidad de un curso en línea con el fin de distribuir la información y conocimiento a lo largo de Kenia y a la vez validar la viabilidad y efecto de los incentivos propuestos para las personas que finalizaran el curso. Los resultados arrojaron que la obtención de insignias y la posibilidad de acceder a una acreditación tiene gran influencia en el desempeño y el interés de los estudiantes en el curso. Para este caso el 75% de los estudiantes demostró mayor interés frente a la certificación y un 25% presentó interés sobre las insignias.

Bajo la misma idea [6](#) plantea una constante incertidumbre respecto al reconocimiento, validación y acreditación de un MOOC, para ello realiza una investigación de validación de la influencia de la acreditación, a partir de una certificación final y/o credenciales digitales y el efecto de estos sobre el comportamiento de los estudiantes. El estudio se realiza a partir de un MOOC de HTML5/CSS introductorio de nivel universitario, que una vez finalizado recibe la percepción y aceptación de los estudiantes mediante entrevistas y encuestas en línea. Los resultados del estudio revelan que existe gran interés por parte de los estudiantes de poder combinar las validaciones no formales como los certificados e insignias digitales con créditos formales.

Por otra parte [18](#) realiza un seguimiento de las actividades de los estudiantes en

un curso MPOC con reconocimiento en el contexto de la Universidad del Cauca. El curso validó dos créditos para cualquier programa de pregrado de la misma institución, siempre y cuando el estudiante culminará el curso aprobando con una nota de 3 o superior en una escala de calificación de 0 a 5. En este escenario no se consideró una actividad de evaluación supervisada ni un seguimiento del comportamiento de los estudiantes con las evaluaciones, por lo que se manifestó incertidumbre por parte de los administradores en la posibilidad de fraude, dado que el desarrollo del curso dentro de la misma universidad podía dar lugar a que los estudiantes se compartieran las respuestas.

### 3.3.2 Deshonestidad académica

#### 3.3.2.1 Examen Supervisado

Bajo esta perspectiva los trabajos destacan dos tipos de categorías, la supervisión personal y la supervisión automática. Para el primer caso [54, 55, 56], identifican una correlación entre la necesidad de supervisión y la incapacidad de las instituciones de abordarla y más cuando la actividad de evaluación es en línea. Los trabajos presentan algunas alternativas de autenticación y supervisión en las que se propone exámenes presenciales supervisados personalmente, es decir el estudiante viaja hasta un centro de pruebas, lo que limita la participación de estudiantes remotos. Otro caso de supervisión se realiza a partir del uso de cámaras web, sin embargo este proceso tiene costos elevados de las instituciones que prestan el servicio y por otra parte requieren de una revisión manual por parte de un evaluador, dando lugar a que el proceso no sea escalable en el entorno MOOC. Finalmente bajo la misma perspectiva se propone un caso de supervisión biométrica, el trabajo realiza una comparación entre la imagen de perfil y la de la tarjeta de identificación, en este caso se presenta una constante preocupación por la autenticación de identidad por lo que el método combina autenticación biométrica con la revisión manual de un examinador, lo que lo hace nuevamente no escalable al entorno MOOC.

Otros trabajos como [57, 58, 59], enfocan su esfuerzo en un tipo de supervisión automática que ya no requiera de un personal de apoyo en la actividad de evaluación

del estudiante. En [57], se propone el uso de SMOWL, una herramienta que utiliza un modelo matemático de identificación de rostros. La validación de su desempeño es buena, sin embargo es una herramienta de pago. Para el caso de [58], presenta un modelo flexible de autenticación del estudiante en el examen en línea, la propuesta del modelo es bastante flexible al tipo de curso y estudiante, no obstante el caso de estudio es hipotético, por lo que se debe validar formalmente en el futuro. Con la misma intención [59], desarrolla un modelo para respaldar la integridad de los exámenes en línea, para ello luego de realizar una búsqueda sistemática de la tecnología actual y los sistemas biométricos utilizados para la supervisión, propone reconocimiento facial, vídeo vigilancia para la supervisión del estudiante y software para limitar la funcionalidad de la computadora.

En [60] se propone una monitorización automática usando dos cámaras y un micrófono. La Webcam de escritorio y la Wearcam que es una cámara que porta el estudiante. Por otra parte se monitorea el audio, lo que da lugar a un campo de visión muy amplio. A partir de ello implementa un algoritmo híbrido que utiliza un clasificador que permite detectar el comportamiento de fraude. No obstante el sistema requiere de elementos externos con los que quizá no cuente el estudiante o que puede afectar su buen comportamiento en el desarrollo del examen.

Por su parte [61] propone un mecanismo multimodal que permite detectar los puntos característicos del rostro del estudiante y a partir del ángulo de guiño de la cabeza detectar malas prácticas. Este trabajo aún debe validarse puesto que con la posición de la cabeza aún no es posible detectar por completo el comportamiento de fraude. Bajo el mismo objetivo [62] desarrolla un preprocesamiento de vídeo, su trabajo detecta el comportamiento anormal y a partir de un análisis de contenido obtiene un resumen de vídeo del examen en línea, de forma que facilite al examinador la tarea de supervisión.

### 3.3.2.2 Análisis de aprendizaje

En [63, 64, 65, 66] proponen el descubrimiento de conductas inapropiadas a partir del análisis de datos de la interacción del estudiante con la plataforma. En [63], Coursera desarrolla un mecanismos de reconocimiento basados en la biométrica de

escritura, este mecanismo aprende la forma en como una persona hace su firma y de esta manera logra reconocer si la persona que escribe es la misma del perfil de la cuenta, no obstante estos mecanismos basados en la biométrica resultan poco fiables y costosos de implementar, además no hacen como tal un seguimiento del estudiante en un curso, en su artículo dicen que para acceder a este servicio se debe pagar entre \$30 y \$100 dolares por cada curso y en beneficio se puede obtener un certificado verificado que demuestra que cumplió con los criterios de aprobación de su riguroso curso en línea. Sin embargo no muestran los resultados de cuantos usuarios han usado el servicio.

En [64] se desarrolla un esfuerzo para detectar CAMEO (Copying Answers using Multiple Existences Online) es decir, detectar una tendencia para hacer fraude a partir de la creación de múltiples cuentas, donde una se denomina recolectora, en la cual se recoge las respuestas correctas a partir de ensayo y error, mientras en una cuenta maestra se envían corregidas. En este trabajo un algoritmo identifica y etiqueta los intentos de esta actividad (CAMEO) mediante la IP de los envíos. Entre los resultados obtenidos se estima que en 1.9 millones de participantes en 115 MOOC de dos universidades, el 1.3% de los certificados se obtuvieron con CAMEO y que estos usuarios tienen más probabilidades de ser jóvenes, hombres y cuentas internacionales.

Por otra parte [65], lo hace a partir de un algoritmo de *Machine Learning* que no requiere de la IP de envío. Este trabajo obtuvo como resultado un nuevo enfoque de detección de CAMEO basado en aprendizaje automático. Si bien se resalta un modelo muy poderoso con eficiencia del 37%, tiene un problema de estimación que debe mejorarse y que solo funciona cuando las evaluaciones son de respuesta cerrada, en las que es posible obtener la respuesta correcta. sin embargo el detector no funciona en un contexto de evaluación por pares de ensayo. El autor plantea la necesidad de evaluar el modelo en diferentes cursos, pues este se probó en un MOOC con un tamaño reducido y propone segmentar el modelo por características para identificar cuales de ellas logran ser más importantes. Aunque la técnica CAMEO usada por los estudiantes para hacer fraude se puede controlar en ambientes más reducidos como los SPOC o MPOC, aún se presentan otro tipo comportamientos inapropiados que hay que categorizar y analizar con el fin de

detectarlos y dificultarlos a partir de la interacción con la plataforma y el dispositivo con el que accede, para lo cual el autor plantea el uso de Cookies<sup>1</sup> como una alternativa.

En cuanto a [66], enfoca su esfuerzo en detectar lo que ha denominado los “vínculos de participación invisibles”, es decir un método para detectar los vínculos de colaboración entre estudiantes. Esto mediante el estudio del tiempo de los envíos de respuesta de un examen. Este trabajo permitió verificar que la mayoría de los estudiantes se agruparon en parejas, aunque también se detectó grupos más grandes. El autor manifiesta que es posible que los grupos se formen con el propósito de realizar fraude, además agrega que se requiere más trabajo en el futuro para buscar estrategias y adoptarlas para tomar medidas en las cuentas de los estudiantes y resalta la gran importancia en descubrir las asociaciones entre las cuentas de los usuarios en los que existe una colaboración con un fin no ético.

En la revisión de estos trabajos es posible evidenciar que se requieren trabajos futuros a la hora de caracterizar los comportamientos inapropiados en el ámbito de la deshonestidad académica a partir del análisis de datos de la interacción del estudiante con la plataforma.

### 3.3.3 Seguimiento al estudiante

#### 3.3.3.1 Recolección de datos

Existen trabajos relacionados con la recolección de datos provenientes de las plataformas MOOC, es decir, muestran herramientas con las cuales es posible obtener los datos de las interacciones de los participantes del curso como los siguientes:

En el caso de [11] propone secciones recomendadas de la plataforma donde se pueden conseguir datos y qué hacer con ellos, da algunas técnicas de procesamiento de datos, sin embargo no muestra cómo hacerlo. Plantea un enfoque a partir de métodos y

---

<sup>1</sup>Archivos creados por los sitios web y almacenados en el navegador, (en este caso creados por las plataformas MOOC) y es donde se guarda información del usuario cuando visita la plataforma.

herramientas de Big data<sup>2</sup> sobre los datos de diferentes cursos dispuestos sobre la plataforma Canvas Network, implementa estrategias de visualización de datos como resultado de un análisis de regresión lineal. Su objetivo radica en que a partir de este análisis se puede desarrollar un mecanismo de sugerencia personalizado. Por otra parte su enfoque no se relaciona con la deshonestidad académica en la actividad de evaluación en un MOOC y el desarrollo no es replicable.

Por otra parte en [68], profesores de la Universidad de Szeged, han enfocado su trabajo a la implementación de una plataforma basada en MOODLE. Su trabajo se centra en la creación de un sistema con algoritmos de registro al curso. Con esto se busca realizar ajustes pedagógicos en primera medida, para luego a partir de la implementación de un algoritmo robusto de aprendizaje automático (ML, Machine Learning), conseguir tareas de adaptación de contenido y notificación de información importante al instructor.

Desde otra perspectiva [69], aporta una revisión sistemática desde 2011 a 2016 en la que se busca identificar los métodos de recolección de datos en la realimentación de los MOOC. Los resultados arrojaron que los métodos de recolección utilizados para estudios cuantitativos incluían, encuestas, entrevistas y análisis de datos de actividad en el MOOC, como el análisis clickStream.

En el trabajo desarrollado en [70], se hace énfasis en el importante trabajo de analítica de datos en las plataformas. Para ello se ha desarrollado un esfuerzo para capturar información de la interacción de los estudiantes con la plataforma y las distintas actividades a partir de las Url de navegación dentro de la plataforma Coursera. Estas Urls contienen información de qué tipo de actividad realiza el estudiante, si ingresa a los contenidos del curso, la interacción social, la interacción con la evaluación del curso, o si ve un vídeo, entre otros. El objetivo de este trabajo es analizar los patrones obtenidos de la navegación de los Urls, más no el comportamiento de los estudiantes con el curso. Por lo que requiere que nuevas investigaciones apliquen tareas de analítica de datos a partir de estos resultados, obtenidos sobre la plataforma Coursera.

---

<sup>2</sup>Información de gran volumen, alta velocidad y gran variedad que exige formas rentables e innovadoras de procesamiento de información para mejorar la comprensión de la misma [67]



### 3.3.3.2 Visualización

En la actualidad existen gran cantidad de datos generados día a día, los vastos volúmenes de esta información han dificultado tanto su interpretación como su análisis. Para ello se ha hecho importantes estrategias de visualización y analítica de datos para ayudar a explorar y extraer información del amplio conjunto de estos y poder responder preguntas a partir de la información dispuesta de una forma más ejemplificada a través de componentes visuales. En el contexto de los MOOC se debe tratar con una gran cantidad de datos, los cuales han promovido enfoques de analítica y visualización [71].

Bajo este contexto [23], plantea una herramienta de visualización llamada MoocViz, la herramienta se fundamenta en técnicas de visualización de datos que permite una vista de información para el análisis del comportamiento del estudiante en el aprendizaje MOOC. Esto a partir de datos en bruto obtenidos de cursos ya desarrollados sobre plataformas como Coursera y Edx. La herramienta permite el análisis del flujo de clicks sobre los vídeos del curso, análisis de foros, entre otros. Los resultados se clasifican a partir de diferentes colores, gráficos circulares de diferente tamaño y líneas de transición. Sin embargo trabaja con los datos de foros y flujos de clicks en vídeo como elementos aislados, y no permite visualizar el comportamiento del estudiante en el ámbito evaluativo y la deshonestidad académica.

En cuanto a la visualización de datos que permitan caracterizar el comportamiento de aprendizaje de los estudiantes en un MOOC se han encontrado trabajos que se basan en el foro del curso, debido a que es uno de los principales componentes de interacción social. Entre estos trabajos se encuentran: Message Lens [22], que permite un análisis visual de los datos del foro, como lo es, saber el nivel de similitud y prevalencia de un tema, cuáles son las palabras más usadas, tener una representación gráfica de la red de interacción entre estudiantes que se forma a medida que se interactúa con el foro. Para ello plantea un modelo de vista grupal, global e individual desde los enfoques de temas de discusión, actitud del estudiante y comunicación entre alumnos.

Con un enfoque semejante VisMooc [24] estudia la cantidad de hilos que se genera a

partir de una pregunta o respuesta en un foro. Por su parte [25] Iforum, desarrolla un sistema de análisis visual que permite explorar los datos del foro desde 3 perspectivas de interacción: las publicaciones, la interacción de los usuarios y los subprocesos de discusión. Promueve además una herramienta llamada "Thread River" que puede ilustrar información de hilos de conversación prolongados.

Por otra parte y con un objetivo más específico en cuanto al análisis de deserción, [26] muestra una forma de representar gráficamente la clasificación de los estudiantes en dos categorías diferenciadas por colores, en la cual se muestran de color naranja los estudiantes con mayor probabilidad de tener éxito en el curso y de un color azul aquellos que no tienen probabilidad, esto a partir de una estrategia de aprendizaje automático y tomando como referencia los datos recolectados en el foro de un curso MOOC de Iversity.

En el caso de [27] T-Cal, se visualiza el tiempo en que se conectó un estudiante, la hora de entrada y hora de salida, presentan gráficos de los estudiantes que estuvieron conectados al mismo tiempo y cuál fue el tiempo de estadía en la plataforma. Toda esta información para la clasificación es basada en los foros de discusión.

Otra técnica usada para saber y analizar el comportamiento de un estudiante de manera visual, es trabajando a partir del flujo de clics, para este caso PeakVizor [28], recolectó información sobre los clics que se hacen sobre los vídeos del curso; con esto estudiaron y representaron gráficamente en que partes del vídeo existía mayor cantidad o picos en la secuencia de clics. Esto arrojó como resultado el descubrimiento de los temas con mayor complejidad para aprenderlos. El trabajo desarrolla 3 vistas en las que contempla una vista general con glifos<sup>3</sup> para visualizar estadísticas importantes sobre los picos detectados, una vista con la distribución temporal de las secuencias y una última que relaciona la correlación existente entre los diferentes grupos de estudiantes y los picos de secuencia de clics en los vídeos.

Basado en la misma técnica, Viseq [72] es una herramienta de análisis visual en la que se pretende no perder la secuencia de las actividades del estudiante. Su objetivo radica en estudiar las actuaciones y comportamientos del estudiante a partir del análisis de patrones de secuencia obtenidos de los datos almacenados

---

<sup>3</sup>Representación gráfica o simbólica de información.

en la plataforma. Plantea cuatro vistas en las que espera reconocer el tipo de grupo de estudiantes, mostrar patrones de secuencia generales de un grupo específico, visualizar transiciones entre las secuencias y una vista que permita la comparación individual (personal) con una secuencia general. Sin embargo ambos trabajos dejan una brecha para la exploración de objetivos más específicos de visualización a partir de la secuencia de actividades y una implementación en grupos de estudiantes más grandes, por lo tanto no desarrollan un análisis visual en el contexto de la deshonestidad académica.

En el caso de [73] se crea un entorno de visualización más amplio: es una solución interactiva para visualizar el progreso del aprendizaje de los estudiantes, contempla 4 vistas, una vista de distribución demográfica para mostrar el lugar donde están tomando curso, una vista del calendario de actividad, para hacer un seguimiento de la actividad de los estudiantes a lo largo de las semanas, una vista de distribución del progreso para obtener una visión general del progreso de aprendizaje de los estudiantes en una semana en particular y la vista de huellas personales para saber sobre el progreso de aprendizaje de un estudiante a través de semanas. Si bien el trabajo desarrolla una herramienta interactiva, requiere de muchos recursos web que a gran escala presenta dificultad, y su enfoque de desarrollo no contempla visualización de datos para explorar una posible incertidumbre de fraude.

# Capítulo 4

## Mecanismo de visualización de datos

En este capítulo se muestra la propuesta que dio continuidad al trabajo previo desarrollado por el Mag. Daniel Jaramillo Morillo. En este trabajo se buscó expandir ese conocimiento y mejorar la visualización de los datos para permitir la generación de información valiosa de los estudiantes y mostrar su comportamiento a los administradores y docentes del curso, esto enfocado en los comportamientos sospechosos de fraude dentro de los cursos MOOC [74, 3]. Aquí se presenta el módulo de recolección de datos, la plataforma OpenEdx y la instancia de la misma implementada en la Universidad del Cauca llamada “SELENE UNICAUCA”. Finalmente se presenta la arquitectura general del mecanismo de visualización denominado “SELENEVIS”.

### 4.1 Recolección de datos

Para llegar al mecanismo de visualización se ha planteado un proceso de recolección de datos, los cuales son seleccionados y categorizados a partir de la información que brinda la plataforma de OpenEdx y sus respectivas instancias, que en este trabajo corresponde a “SELENE UNICAUCA”.

En esta subsección se presentan las principales características de la plataforma Selene, la forma en como organiza los datos y además, se abarca brevemente las características de sus bases de datos y el archivo “tracking.log”, archivo que permite acceder a la información de la interacción de los estudiantes con la plataforma en forma de eventos [74].

### 4.1.1 Plataforma de aprendizaje

#### 4.1.1.1 Open Edx

Es una plataforma de código abierto creada por la asociación de dos instituciones de alto prestigio como son la Universidad de Harvard y el grupo de investigación del MIT (Instituto Tecnológico de Massachusetts). Edx ofrece gran variedad de cursos MOOC a través de su portal web edx.org. También posibilita a instituciones y universidades la creación de sus propias instancias de Open Edx para la distribución y despliegue de cursos propios, para lo cual permite la instalación de un servidor de producción que facilita la organización y disposición de su material, la gestión de la plataforma en diferentes niveles, permitiendo a las instituciones desarrollar herramientas de aprendizaje y añadir nuevas características a la plataforma [75].



Figura 4.1 Logo Open Edx. Tomada de [2].

#### 4.1.1.2 Selene Unicauca

Selene Unicauca es una iniciativa de la Universidad del Cauca, apoyada por su departamento de Telemática, que busca a partir de la innovación tecnológica y la educación, fomentar la creación y el despliegue de cursos MOOC para la adquisición y desarrollo de habilidades profesionales, conocimientos y nuevas competencias. Permite a estudiantes, universidad y a gremios que quieran asociarse, contar con

las herramientas necesarias para la construcción y despliegue de sus propios cursos MOOC o simplemente la inscripción para poder acceder a los contenidos aprovechando todo el material educativo digital e interactivo desde diferentes dispositivos como móviles, computadoras, tablets entre otros. Esto a través de la dirección web selene.unicauca.edu.co.

Selene Unicauca es una plataforma que está soportada por los proyectos de MOOC-Maker, MOOCMenTES y MOOC-Unicauca. Brinda la posibilidad de desplegar cursos MOOC desde el año 2016 y actualmente es una plataforma que soporta una gran cantidad de cursos a los que los estudiantes y personas de otros gremios pueden acceder solicitando una inscripción a un determinado curso para validar créditos educativos dentro de la Universidad del Cauca en el ámbito de las electivas FISH (Formación Integral, Social y Humana) o para adquirir nuevas habilidades.

El departamento ha desarrollado esfuerzos por fortalecer este tipo de estrategias y modalidad de aprendizaje por lo que ha desarrollado estudios y ha publicado artículos en prestigiosas revistas. En la Actualidad un enfoque está en la validación de créditos educativos que ha traído consigo la incertidumbre de comportamiento con sospecha de fraude [1, 76].

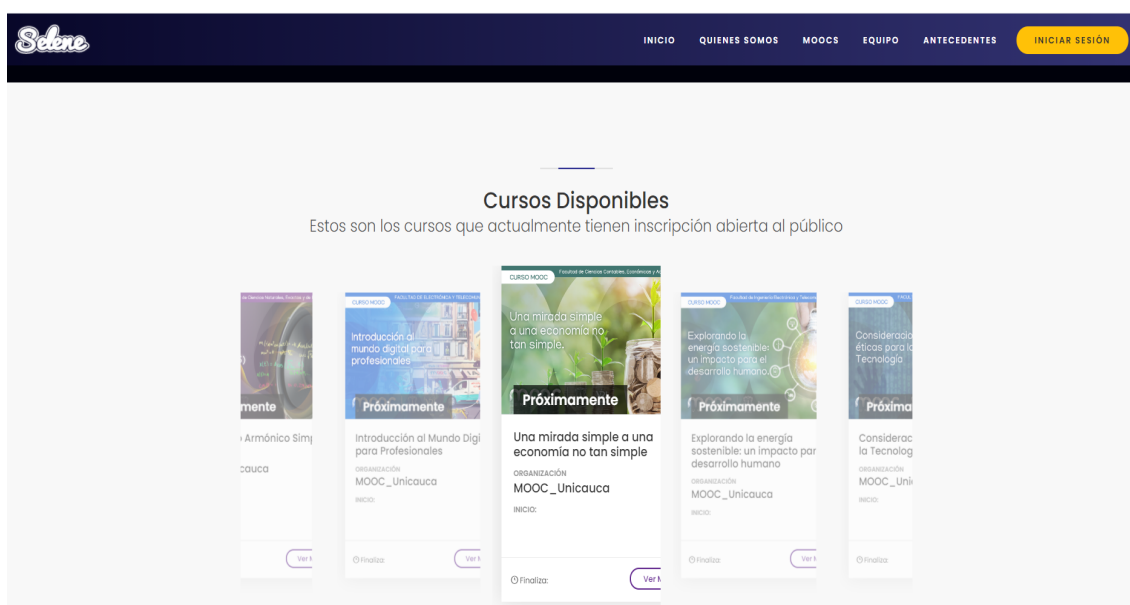


Figura 4.2 Selene Unicauca Inicio.

### 4.1.2 Características Básicas Open Edx

Open Edx es una plataforma de código abierto que puede ser descargada e instalada a nivel de producción a partir de nuevas instancias para ser gestionada y complementada con la generación de nuevas funciones y componentes para la construcción y disposición de nuevos cursos MOOC.

Además ofrece características como la posibilidad de crear diversas formas de evaluación con diferentes configuraciones (autoevaluación, evaluación por pares, evaluación automática y por analítica de aprendizaje, entre otras), crear y gestionar foros de discusión, entornos de simulación interactivos, agregar nuevos módulos a la plataforma a partir de Xblocks, verificar el progreso de las actividades generando informes, incluir material bibliográfico, libros, notas de estudio en formatos como el PDF.

También facilita anexar imágenes, visualizar contenido en vídeo grabado con subtítulos y con la posibilidad de ser indexado sobre el mismo, elegir y gestionar el idioma de la plataforma (subtitulación), programar y configurar calendarios de exámenes, inicio y terminación del curso, anexar y establecer wikis de estudio.

Una vez realizada la implementación de la plataforma, ofrece como características básicas e importantes los siguientes componentes:

- LMS.
- CMS.
- Módulo XBlock.
- ORA2 XBlock.
- Open edX Insights.
- Base de Datos Relacional (MySQL).
- Base de Datos no Relacional (MongoDB).

- Tracking log.

De las cuales se trabajo con el LMS, CMS y tracking.log, las cuales se describen en la sección [4.1.2.1](#), [4.1.2.2](#) y [4.1.2.3](#) respectivamente.

#### 4.1.2.1 Sistema de Gestión de Aprendizaje

Los estudiantes, grupos de cursos e instructores usan el LMS a lo largo del curso y cuenta con diferentes vistas como:

- Vista del docente: en esta vista los docentes pueden publicar cursos, foros, comunicarse con estudiantes, generar informes y administrar el curso a medida que se ejecuta [77](#), [76](#).
- Vista del estudiante: el estudiante puede acceder a su perfil, inscribirse en un curso, buscar cursos y mirar su progreso, acceder a contenidos de un curso en especifico, como vídeos, libros de texto, las diferentes secciones, subsecciones del curso, entre otros [77](#), [76](#).

#### 4.1.2.2 Sistema de Gestión de Contenidos

Esta característica de la plataforma también es llamada “Studio” y es en donde se puede gestionar la estructura de los cursos, es decir, cuantas secciones, subsecciones, exámenes tendrá el mismo. También es posible gestionar el contenido, los cuales se dividen en 4 tipos de componentes que se pueden gestionar como; foros, ejercicios, vídeos e incluso pequeños componentes de Lenguaje de Marcado de Hipertexto (HTML, HyperText Markup Language)<sup>1</sup>.

Es de aclarar que para acceder a este apartado es necesario contar con permisos adicionales y que solo los puede otorgar el administrador de la plataforma.

---

<sup>1</sup>HTML: lenguaje de marcado para la elaboración de páginas web [78](#)



### 4.1.2.3 Tracking.log

En este archivo es donde se guardan cronológicamente todos los registros(logs) de las actividades que realiza un usuario dentro de la plataforma. La ubicación de este archivo esta en “/edx/var/log/tracking/” y este archivo es el centro de información de todo el sistema de detección de sospechosos en los trabajos [18, 76] de los cuales parte en este trabajo para realizar la visualización de los comportamientos con sospecha de fraude.

## 4.1.3 Extracción de datos desde Selene Unicauca

La recolección de datos se realiza teniendo en cuenta las consideraciones de funcionamiento de la plataforma Selene Unicauca y la estructura de los componentes explicada en la sección 4.1.1 y 4.1.2 del presente documento. A continuación se muestra el diagrama general del mecanismo de recolección de datos.

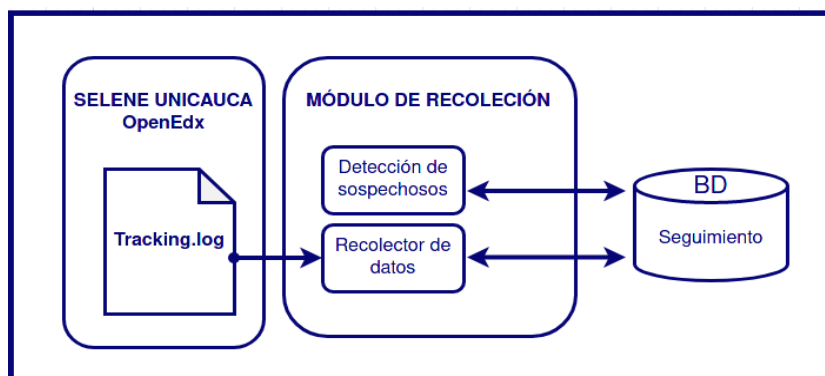


Figura 4.3 Arquitectura para recolección de datos.

La extracción de los datos parte desde el archivo Tracking.log que se encuentra alojado en la plataforma Selene Unicauca. Este archivo guarda el registro de las interacciones con la plataforma por parte de los estudiantes. Se almacenan en forma de eventos y con una estructura de formato de notación de objetos JavaScript (JSON, JavaScript Object Notation). Posterior a esto un módulo llamado Recolector de datos, es el encargado de leer este archivo, sacar una copia, extraer y organizar la información en una estructura adecuada para el mecanismo de seguimiento y

visualización y anexarla a la base de datos "seguimiento.bd". También se tiene un módulo de detección de sospechosos, que permite a partir de la información recolectada en la base de datos, identificar estudiantes sospechosos a partir de la similitud de respuestas y los tiempos de envío de examen, cómo en [1].

## 4.2 Caracterización de datos

La figura 4.4 muestra la estructura de los documentos JSON almacenados en la base de datos "Seguimiento" descrita en 4.1.3, en la cual se observan diferentes campos, los cuales describen un el registro ordenado del archivo "tracking.log". Se hizo un proceso de anonimización de los nombres de los estudiantes para cuidar la integridad de los mismos.

```
{
  "_id" : ObjectId("5f4872f05e738530e80b91ec"),
  "username" : "Esrudiante x",
  "answers" : "Null",
  "course" : "Unicauca+LeanStartUp+2020-I",
  "session" : "93f0b959f716d5991b9b636756d6a0b2",
  "date" : "2020-08-28",
  "unit" : "1",
  "date_time" : "2020-08-28T01:31:32.090349+00:00",
  "name" : "nav_content_click",
  "section" : "5b849ceb504146e9b2e8a9261f6e1537",
  "subsection" : "042a4be8d7e8461893d459c37db09a4a",
  "time" : "01:31:32",
  "page" :
  "/courses/course-v1:Unicauca+LeanStartUp+2020-I/courseware/5b849ceb504146e9b2e8a9261f6e1537/042a4be8d7e8461893d459c37db09a4a/1",
  "seccion" : "Descripcion del curso",
  "sub_seccion" : "Actividades inicio del curso",
  "__v" : 0
}
```

Figura 4.4 Estructura de Registro JSON.

Con el fin de describir el comportamiento del estudiante dentro de la plataforma, se tomó los siguientes campos desde los registros del archivo tracking.log.

- **username:** este campo identifica al estudiante dentro de la plataforma, lo que permite filtrar todos los eventos de un estudiante y hacer el respectivo análisis.

- **course:** es el nombre del curso donde se generó la actividad, el cual permite filtrar los datos por cursos que existan en la plataforma.
- **session:** representa el identificador de la sesión en la que se realizó la actividad. Permite obtener las actividades que se realizan en sesiones diferentes.
- **date:** permite filtrar las actividades de un estudiante por día, ya que representa el día en que se realiza la actividad.
- **name:** este campo describe el nombre del evento que se realiza en la plataforma y uno de los más importantes, pues es el campo que permite conocer la actividad exacta realizada.
- **time:** muestra la hora en la que se realiza una actividad, el cual permite agilizar consultas, por ejemplo, para encontrar las diferencias de tiempos de envío de exámenes, también en algunos casos se tiene en cuenta el campo “date-time”, para consultas en orden cronológico.
- **seccion:** es el identificador de una sección del curso, esto permite conocer cuantos exámenes tiene el estudiante en una determinada sección, ya que se encontró que en algunos casos existen dos registros de un mismo examen.

### 4.2.1 Caracterización de los datos

Para la caracterización de los datos se realiza un proceso de agrupación con el fin de disminuir la cantidad de actividades y quedar con las categorías de “Videos”, “Foros”, “Contenidos”, “Exámenes” y “Sesiones”.

La figura [4.5](#) muestra el resumen del proceso de caracterización de eventos.

### 4.2.2 Construcción de grafos

Para describir el comportamiento del estudiante se propone crear gráficas basadas en grafos, donde los nodos representan las categorías descritas en la tabla [4.5](#) y las aristas el orden con que el estudiante realizó los eventos en la plataforma.

Eventos	Referentes	Categoría
nav_content, nav_content_prev, nav_content_click, nav_content_next, nav_content_tab	Interacciones con menús y material del curso.	Contenidos
stop-video, pause-video, play-video	Interacciones con videos.	Videos
Signin	La pagina principal de la plataforma.	Home
problem-check, problem-graded	Interacciones con Exámenes.	Exámenes
edx.forum.comment.created, edx.forum.response.created, edx.forum.thread.created	Interacciones con los foros del curso.	Foros
session	Identificación de cada sesion de un estudiante.	sesiones

Figura 4.5 Caracterización de eventos.

La figura 4.7 muestra las actividades que se realizan para la creación de grafos. Se inicia con la obtención de los cursos de la plataforma, luego, se obtienen los estudiantes por cada curso, y por cada estudiante se consultan las actividades, las cuales están cronológicamente ordenadas con el fin de replicar el proceso que realizó el estudiante. Paralelo a esto se consultan los días y sesiones por cada estudiante en los cuales se tiene interacción. Luego de esto, se realiza un proceso de caracterización descrito en 4.2.1.

Finalmente se construye la lista de grafos por día o por sesión y se realiza un proceso de filtrado de nodos y aristas iguales, es decir si ya existe una relación del nodo A al nodo B y tiene el título de arista “x”, entonces no se repite la relación, solo se agrega al título de arista un mensaje “y” que representa el siguiente numero en orden cronológico en el que se realiza la actividad, dando como resultado un título “x, y”, lo que permite entender que existe un recorrido adicional al primero tal como se muestra en la figura 4.6, donde se evidencia que la primera interacción se realiza del

nodo A hacia B, luego hay una segunda relación desde el nodo B hacia el nodo A y finalmente como tercer desplazamiento se realiza una interacción del nodo A a B. Por otra parte la estructura de un nodo de grafo se muestra en la figura [4.8](#).

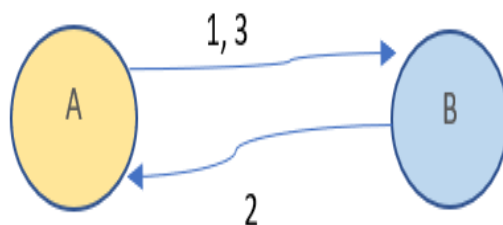


Figura 4.6 Ejemplo control de aristas y nodos.

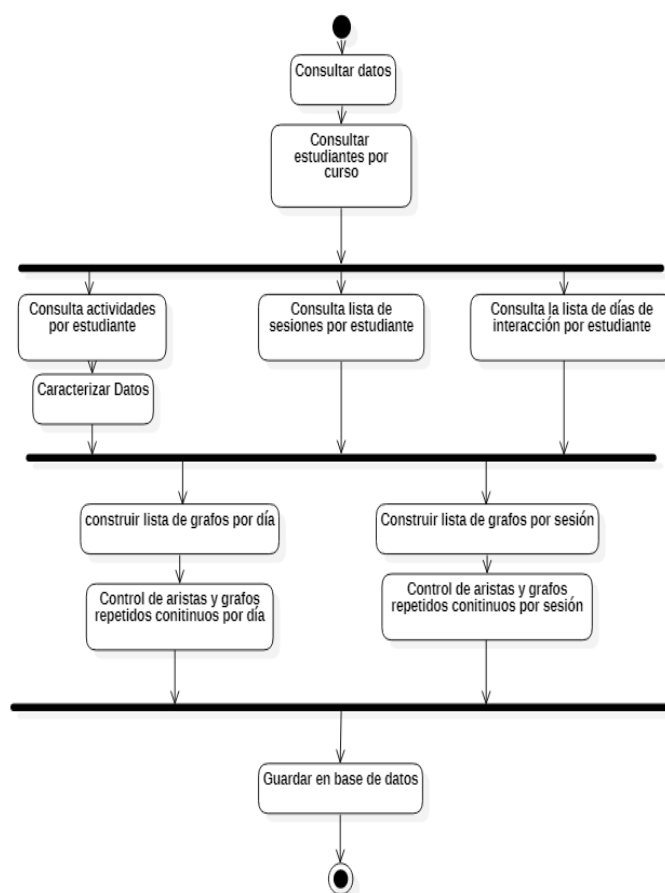


Figura 4.7 Creación de grafos.

```
let edgeOrder = {  
  
  idOrder: `${element.id}${elementnext.id}`,  
  to: elementnext.id,  
  nameto: elementnext.label,  
  from: element.id,  
  namefrom: element.label,  
  label: (countOrder + 1).toString(),  
  visits: countOrder + 1,  
  length: 350  
}
```

Figura 4.8 Estructura de grafos.

En la figura [4.8](#) se observa los campos que conforman los objetos del grafo, donde se tiene un evento llamado “element” y un evento próximo llamado “elementnext”, es decir, el evento actual y el evento siguiente. A continuación se describe los campos de la estructura del objeto.

- **idOrder**: hace referencia a la identificación del objeto.
- **to**: especifica el nodo donde debería conectar.
- **nameto**: indica el nombre del nodo donde debe conectar.
- **from**: identificación del nodo de donde inicia la arista.
- **namefrom**: nombre del evento de donde comienza la arista.
- **label**: es un texto que indica el número de la secuencia sobre la arista.
- **length**: hace referencia al tamaño que tiene el nodo.

#### 4.2.2.1 Presentación de Grafos

Una vez se tiene la Colección "Grafos" en la base de datos de "seguimiento", es posible hacer la petición desde la interfaz de usuario de la aplicación web de SELENEVIS ingresando a la vista de "comparación", los cuales mediante la librería "Vis.js" propia del lenguaje de programación "JavaScript", permite a partir de la estructura de datos mencionada, visualizar el comportamiento de los estudiantes a partir de nodos y sus relaciones, tal cómo se señaló en la sección [4.2.1](#).

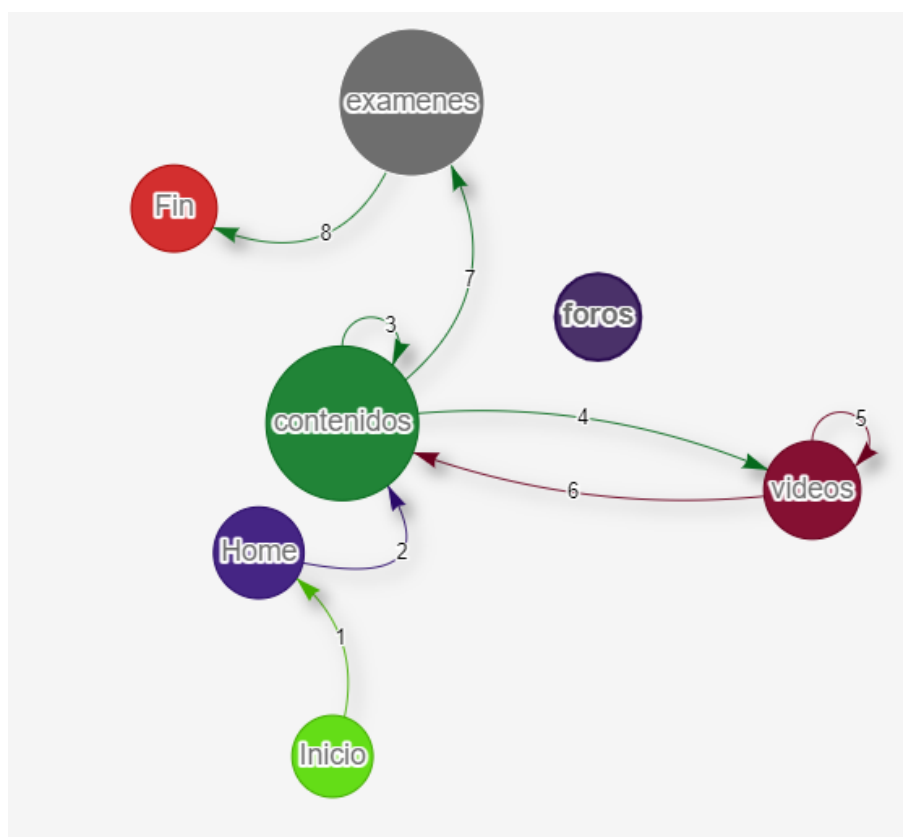


Figura 4.9 Representación visual de grafos.

La figura [4.9](#), permite visualizar de manera cronológica el orden de interacciones del estudiante con las actividades del curso por día o sesión, como se muestra en la sección del capítulo de Resultados [5.1.4](#). Cada uno de los Nodos representa el tipo de actividad caracterizada que corresponden a los señalados en la sección de

caracterización [4.2.1](#). Las aristas, corresponden a las líneas de unión que identifican la dirección de los sucesos. Permiten además de direccionar, enumerar e ir agrupando los diferentes saltos de un nodo al otro a lo largo de un orden cronológico, identificar cuando se queda realizando la misma actividad, como en el caso del nodo de contenido con arista “3”. De igual forma, se puede observar el nodo de actividad con el que el estudiante no tuvo interacción a lo largo del día o a lo largo de una sesión.

## 4.3 Visualización

La aplicación consiste en recolectar datos desde la plataforma Selene Unicauca y visualizarlos. Para ello, se organiza y mejora el formato de los datos recolectados y se guardan en nuevas colecciones dentro de la base de datos “Seguimiento” que es de tipo NoSQL y que usa el motor de MongoDB. Así, esta aplicación permite al docente visualizar el comportamiento del estudiante con base a su interacción con la plataforma de aprendizaje. Se retoma un trabajo previo en donde se busca la detección de los estudiantes con sospecha de fraude, y con base en eso se define una propuesta de visualización [3](#).

### 4.3.1 Arquitectura general

En esta sección se presenta la arquitectura general del mecanismo de visualización y se realiza una descripción de los diferentes módulos y componentes del sistema, los cuales se detallan a través de diferentes perspectivas, a partir de la descomposición del modelo 4+1 Vistas. La figura [4.10](#) muestra el esquema general de la arquitectura donde funciona la aplicación SELENEVIS.

#### 4.3.1.1 Módulo Recolector datos

Este módulo se encarga de leer, crear una copia, recolectar y procesar los datos del archivo tracking.log descrito en la sección [4.1.2.3](#). Lo que hace, es obtener la información desde plataforma y darle un formato adecuado para el seguimiento de



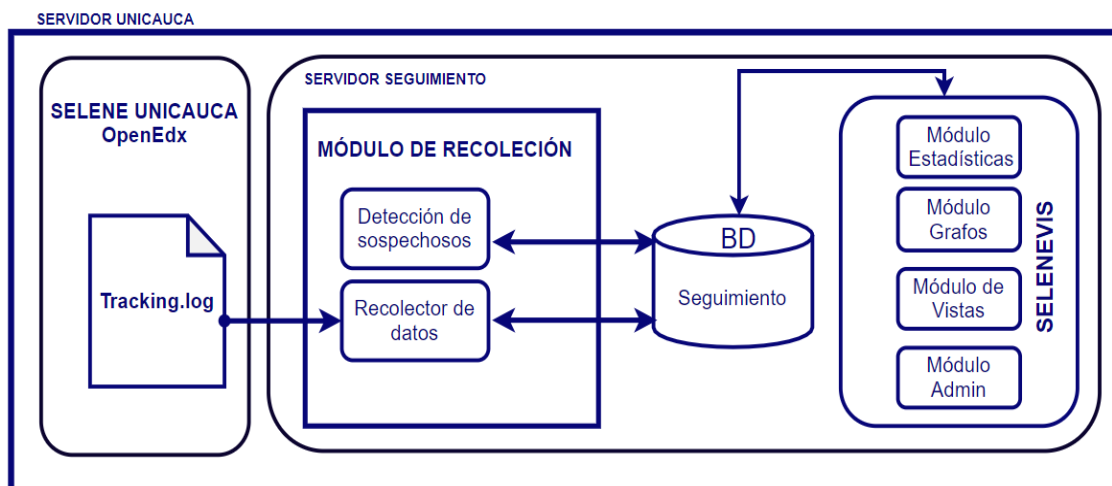


Figura 4.10 Arquitectura General Mecanismo SeleneVis. Adaptada de [3].

las interacciones de los estudiantes en los cursos. Este módulo se retoma de un trabajo previo de la Universidad del Cauca [18].

#### 4.3.1.2 Módulo Detector sospechosos

Este módulo es el encargado de obtener, clasificar y guardar los datos de los estudiantes con sospecha de fraude. Este módulo detecta a los sospechosos de acuerdo a la similitud de estudiantes respecto al porcentaje de coincidencia de respuestas a lo largo de los diferentes exámenes del curso y también, de acuerdo a sus tiempos de envío. Este módulo fue diseñado y desarrollado completamente por el estudiante de Doctorado Daniel Jaramillo Morillo [1].

#### 4.3.1.3 Módulo Estadísticas

Encargado de procesar y guardar las interacciones de los estudiantes alojadas en la base de datos “Seguimiento”. Este módulo permite identificar un resumen estadístico de las interacciones para cada uno de los estudiantes en cuanto a vistas de contenido, interacciones con foros, exámenes, vídeos y el número de sesiones realizadas a lo largo

del curso. Esta información se actualiza en tiempos programados para mejorar el rendimiento de consultas desde la interfaz de la aplicación Web y se guarda en la base de datos "seguimiento" en una colección propia de la aplicación SeneVis llamada "Estadísticas".

#### 4.3.1.4 Módulo Grafos

Se encarga de procesar la información de los estudiantes dispuesta en la colección "Registros" de la base de datos "Seguimiento". Este módulo genera un nuevo formato de documento en la base de datos que establece las conexiones entre las diferentes actividades, permitiendo ver la continuidad en orden cronológico de cómo interactuaron los estudiantes con el curso. Además permite clasificar estas conexiones por día o por sesiones en las que los estudiantes tiene interacción con las actividades del curso.

#### 4.3.1.5 Módulo Admin

Es el responsable de permitir todas las operaciones con los registros de los docentes, es decir, registrar, eliminar, modificar y obtener la información de los mismos dentro de la plataforma. Por otra parte permite la creación de nuevos usuarios administradores y la gestión de los mismos dentro de la aplicación Web.

#### 4.3.1.6 Módulo Vistas

Este módulo hace referencia a todas las vistas de la aplicación, es decir es el encargado de servir las diferentes interfaces web que existen en la aplicación, las cuales fueron diseñadas con el framework de Angular<sup>2</sup>. Este módulo permite presentar bajo diferentes niveles de abstracción la información de estadística del curso y de los estudiantes. Para ello dispone de la vista de Estadísticas, Vista de Sospechosos,

---

<sup>2</sup>Entorno de trabajo o framework de código abierto que facilita la creación de aplicaciones web

vista de Comparación y la vista de Administrador que se describen con detalle en el capítulo 5 “Resultados”.

## 4.4 Descripción arquitectura SELENEVIS

La arquitectura de la aplicación se describe a través del modelo “4+1 vistas” diseñado por el profesor Philippe Kruchten y encaja con el estándar “IEEE 1471-2000” (Recommended Practice for Architecture Description of Software-Intensive Systems) [80]. La figura 4.11 muestra las diferentes vistas que se tienen para la descripción de la arquitectura utilizando este modelo.

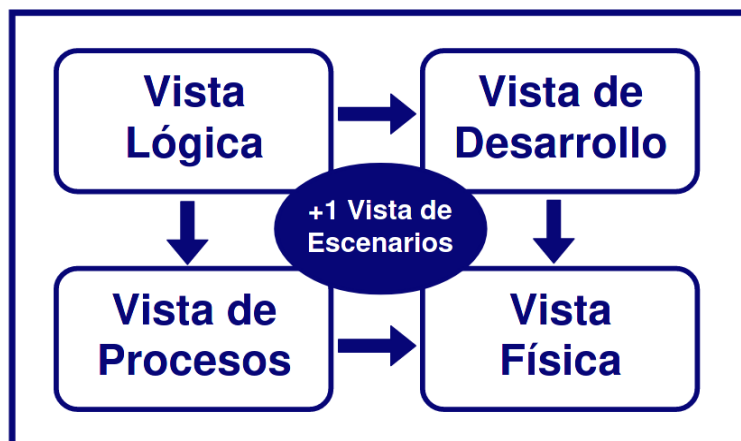


Figura 4.11 Modelo de 4+1 vistas del Proceso Unificado de Desarrollo (Rational Unified Process, RUP)..

### 4.4.1 Vista de escenarios

La vista de escenarios es la que relaciona todas las demás vistas y se describe a través de diagramas de caso de uso, además es obligatoria cuando se usa el modelo 4+1, ya que todas las demás vistas se derivan de ella [76, 80]. A continuación se listan los requerimientos funcionales de la aplicación seguido de la vista de escenarios.

- Iniciar sesión.
- Cerrar sesión.
- Visualizar estadísticas de un curso.
- Visualizar estadísticas de un estudiante.
- Visualizar grafos de comportamiento.
- Comparar estadísticas de estudiantes.
- Comparar estudiantes sospechosos.
- Gestionar docentes.
- Visualizar exámenes de un curso.
- Visualizar estudiantes sospechosos.
- Comparar estudiantes sospechosos.
- Gestionar docentes.

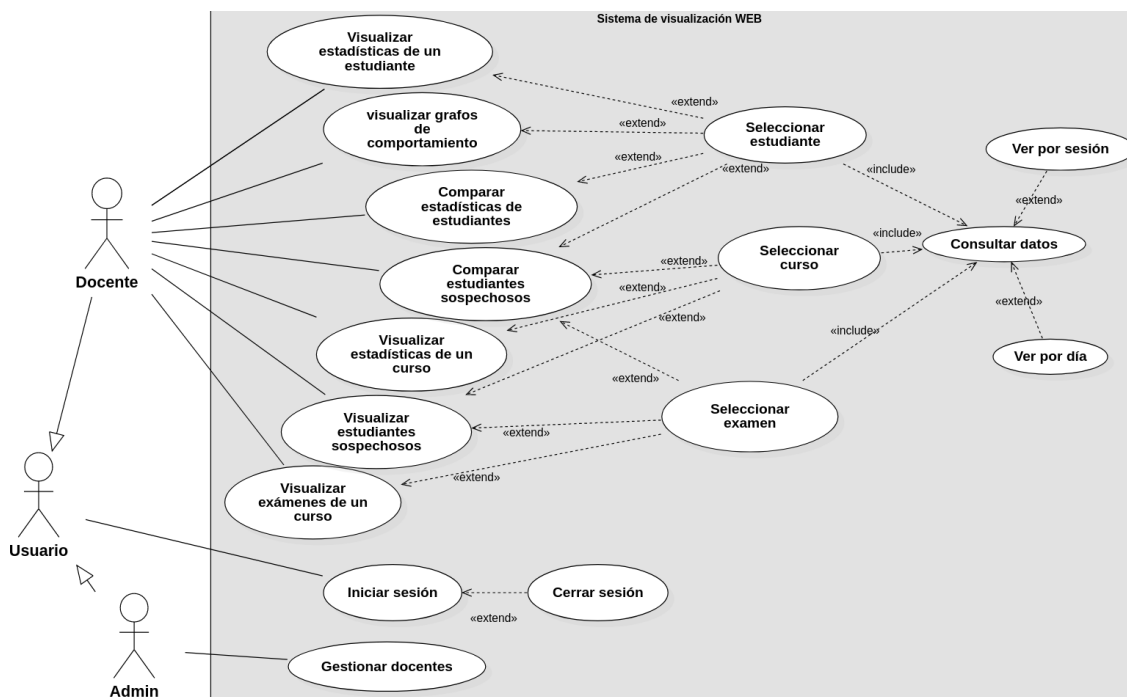


Figura 4.12 Diagrama de casos de uso.

La herramienta debe permitir a los docentes o administradores visualizar las estadísticas de un estudiante o de un curso, que para este trabajo hace referencia a las interacciones de los estudiantes con los contenidos del curso. También poder comparar dichas estadísticas entre estudiantes, para ello se debe seleccionar un estudiante y un curso. Estas estadísticas pueden ser vistas por días o por sesiones en las que interactuó el estudiante.

Debe permitir la visualización de los diferentes exámenes que existen en un curso

e incluso los estudiantes sospechosos del curso en cuestión y los sospechosos que existen en el examen en particular, para ello se debe seleccionar el examen del cual se quiere dicha información.

Además debe permitir la visualización de grafos, los cuales describen el comportamiento de los estudiantes a partir de las interacciones del curso, estos pueden ser visibles por días o por sesiones en las que interactuó el estudiante.

#### 4.4.2 Vista lógica

La vista lógica es un apoyo a los requerimientos funcionales y lo que busca representar son servicios que se presentan al usuario final. La representación se hace a través de clases de objetos, se aplican los principios de abstracción, encapsulamiento y herencia [80].

La figura 4.13 presenta el diagrama general del sistema de Visualización, basado en los trabajos [18, 76], y la figura 4.14 detalla el diagrama de clases específicamente de SELENEVIS.

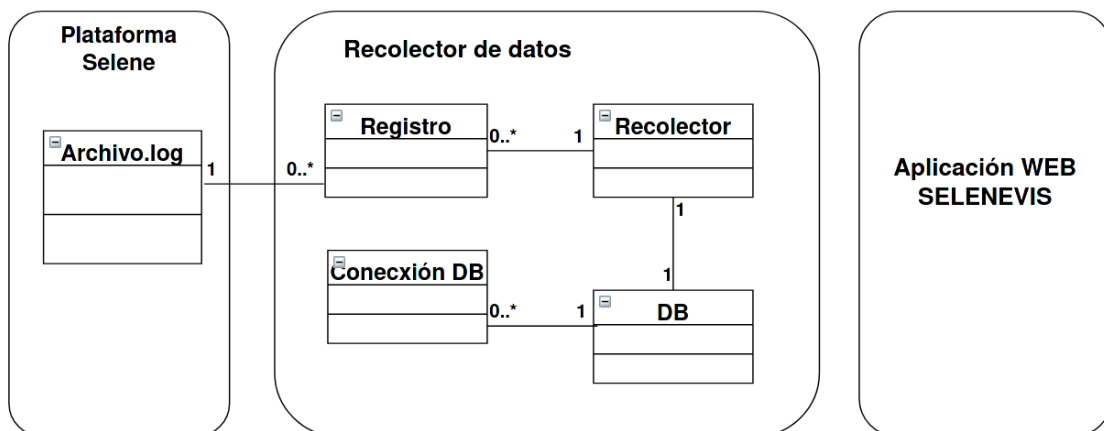


Figura 4.13 Diagrama de clases general.

La clase `Archivo.log` hace referencia al archivo `tracking.log`, donde se encuentran todas las interacciones con la plataforma. La clase `registro` es la estructura que se

le da a los datos. El recolector es el encargado de procesar y guardar los registros en la base de datos [3].

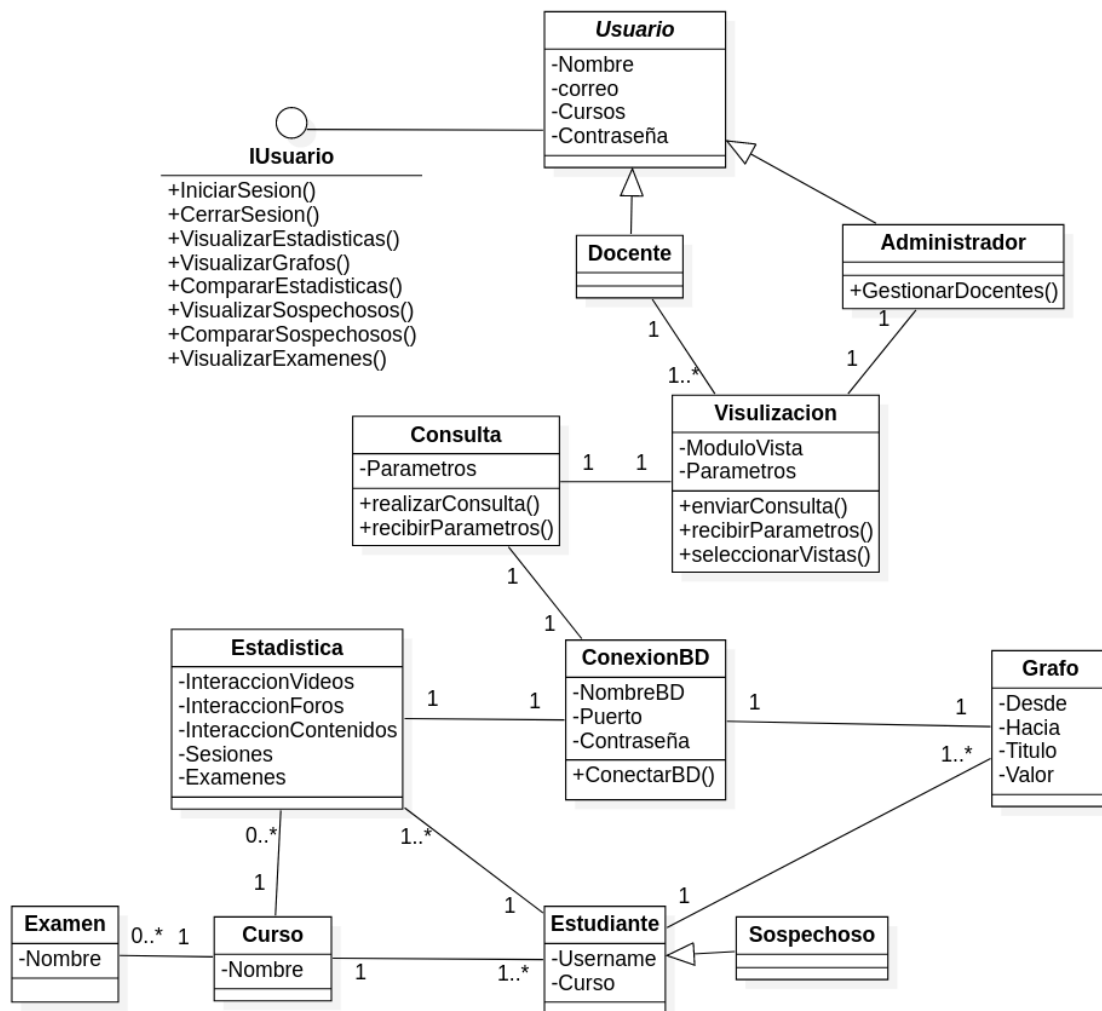


Figura 4.14 Diagrama de clases SELENEVIS.

En cuanto a SELENEVIS, en la figura 4.14 se puede observar la clase abstracta “usuario” que implementa la interfaz “IUsuario” que posee la mayoría de los métodos generales que heredaran la clase “Docente” y la clase “Administrador”, las cuales están asociadas a la clase “Visualización”, que permitirá visualizar las diferentes vistas en la aplicación y que esta asociada a la clase “Consulta”, encargada de realizar las consultas usando la clase “ConexionBD”, que dependiendo de los parámetros de

consulta, se establece en “Estadística” o “Grafo”, las cuales están relacionadas con los cursos y con los estudiantes.

### 4.4.3 Vista de procesos

En esta vista se observan los diferentes procesos y cómo éstos se comunican. Para ello se describen los flujos de trabajo del negocio y del sistema mediante el uso de diagramas de actividades que están relacionados con los casos de uso del sistema mencionados. Para este caso particular, se dividen en tres módulos, el módulo de recolección, el módulo de detección de sospechosos y el módulo de visualización.

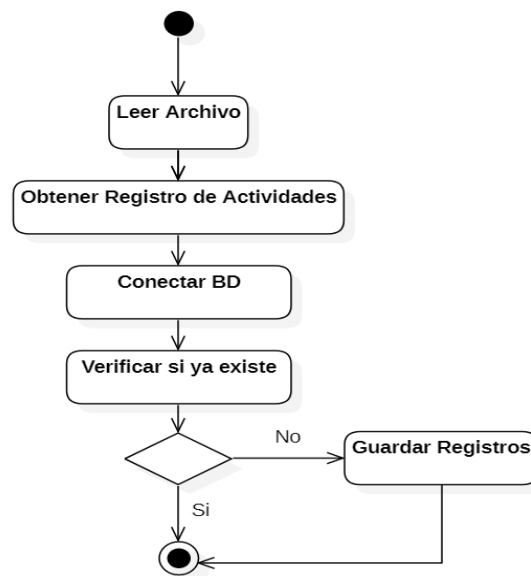


Figura 4.15 Diagrama de Proceso-Recolección De Datos.

La Figura [4.15](#) muestra el proceso de Recolección de datos que inicia con la lectura del archivo de extensión .log. Esto permite obtener los registros de las actividades de los estudiantes con la plataforma y los almacena en memoria en formato JSON. Posteriormente se realiza un filtrado de estos registros de acuerdo a la información acorde a las necesidades del sistema, y una vez completado este paso, se realiza una conexión a la base de datos y se valida si el registro ya existe, de no estar registrado, se añade y guarda en la base de datos como un nuevo elemento.



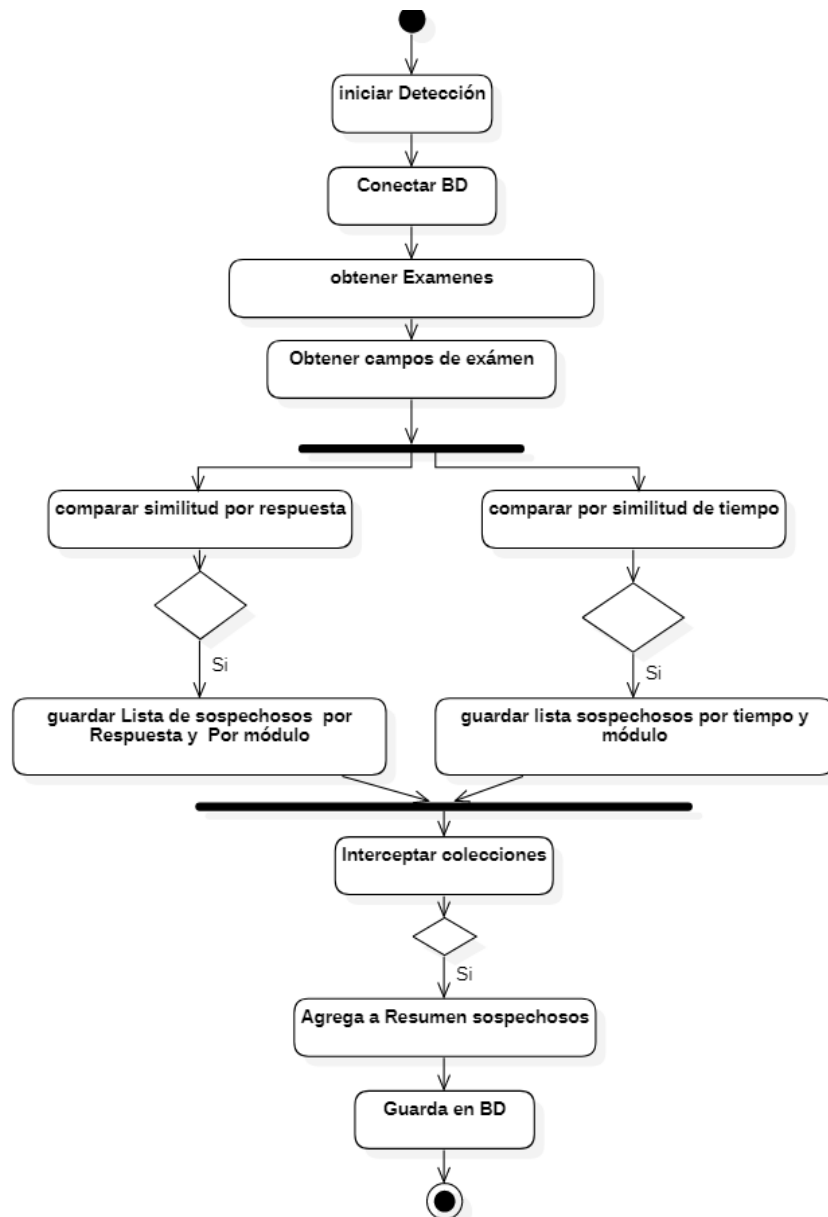


Figura 4.16 Diagrama de Proceso-Detección de Sospechosos.

La figura [4.16](#), permite visualizar el diagrama de actividades, en el proceso de detección de estudiantes sospechosos de fraude realizado en el trabajo de [1](#). El proceso inicia con una petición que requiere de una conexión a la base de datos para la obtención de los registros de los diferentes módulos de examen para todos los estudiantes. Simultáneamente se compara la similitud de respuesta para todos los

estudiantes en los diferentes módulos y de igual forma la similitud de las respuestas y diferencias en tiempos de envíos. Se procede a interceptar las colecciones de los dos casos para generar un resumen de sospechosos, que posteriormente se almacenan en la colección de sospechosos en la Base de datos "Seguimiento" de la plataforma Selene Unicauca.

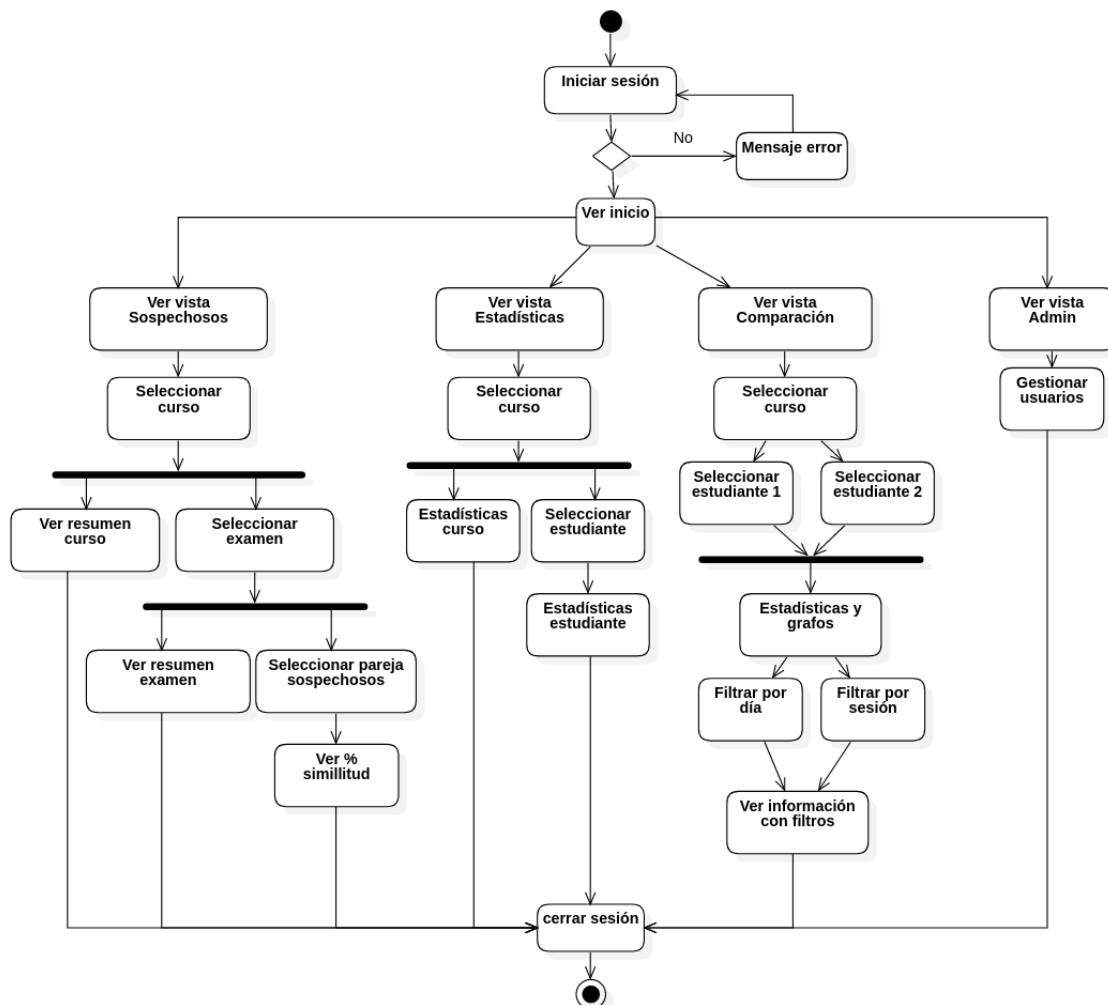


Figura 4.17 Diagrama de Proceso-Módulo de Visualización.

La figura [4.17](#) muestra el flujo del proceso a partir del diagrama de actividades. Este proceso requiere del inicio de sesión, si este es correcto, entonces se habilita la pantalla de inicio dónde se puede seleccionar las diferentes vistas de la aplicación (Vista de Sospechosos, Vista de Estadísticas, Vista de Comparación y Vista de

Administrador). Cada flujo del proceso para las vistas se especifica en el diagrama de actividades y su funcionamiento se describe a detalle en el capítulo 5 “Resultados” de este documento.

#### 4.4.4 Vista de Desarrollo

La Vista de Desarrollo permite observar las conexiones de los diferentes componentes del sistema y sus relaciones. Esta Vista se ocupa de la gestión del software, por lo que se ha creado un diagrama de componentes que incluye los diferentes módulos de la aplicación Web “SeleneVis”.

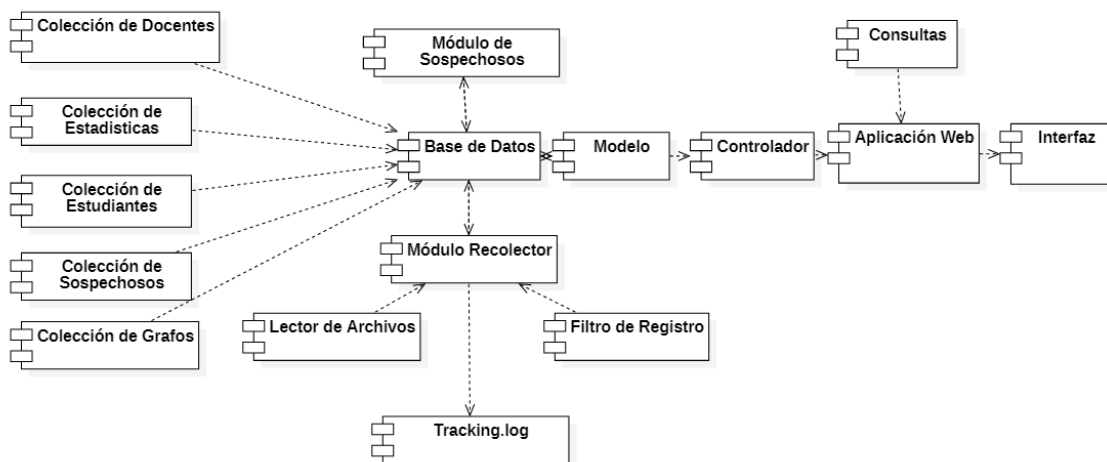


Figura 4.18 Diagrama de Desarrollo Mecanismo de Visualización.

#### 4.4.5 Vista física

La vista física toma en cuenta los requerimientos no funcionales del sistema tales como la disponibilidad, confiabilidad (tolerancia a fallas), rendimiento y escalabilidad. Las conexiones que conforman el sistema incluye también los servicios internos.

En la figura 4.19 se muestra la configuración para SELENEVIS, en la cual se aprecia que dentro del servidor de “Unicauca” existen máquinas virtuales, en las

cuales estarán alojadas las aplicaciones de “SELENEVIS”, la base de datos llamada “Seguimiento” y la plataforma “Selene Unicauca”.

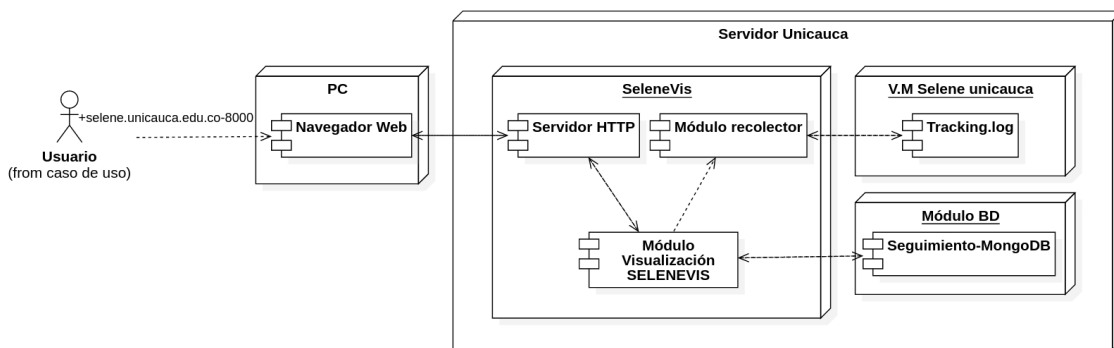


Figura 4.19 Vista física.

## 4.4.6 Implementación

### 4.4.6.1 Repositorio

Con el fin de tener un mayor control de las versiones del desarrollo de la aplicación, un desarrollo colaborativo entre diferentes desarrolladores (en este caso dos) y dar la posibilidad de que en trabajos futuros mas desarrolladores se puedan unir al proyecto, se crea un repositorio <sup>3</sup> de GitLab que es un sistema basado en Git <sup>4</sup> que permite la administración de código fuente a través de herramientas como “Continuous Integration” (CI), “Continuous Delivery” (CD) y Runners, que son servicios que permiten enlazar el repositorio del código fuente con el servidor o maquina virtual donde correrá la aplicación y hacer actualizaciones de manera continua a la aplicación web, mas conocido como DeVOps (Development Operations) <sup>[81]</sup>.

<sup>3</sup>Lugar de almacenamiento centralizado donde se administra código fuente a través de un software como Git

<sup>4</sup>Software creado por Linus Torvalds, para el control de versiones de aplicaciones software.

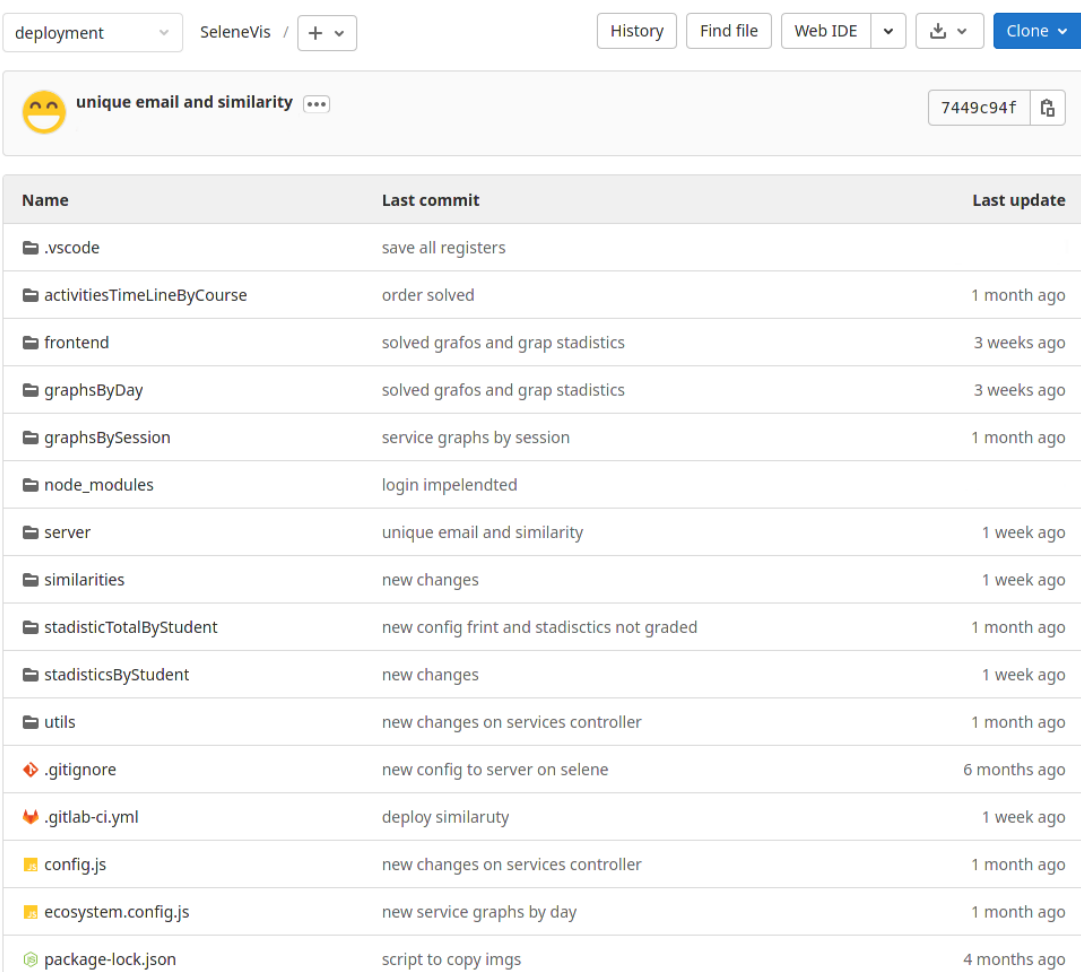
### 4.4.6.2 Servidor

Para la implementación de la aplicación web se tiene una maquina virtual con las siguientes características en un servidor de la Universidad del Cauca:

- 8Gb de memoria RAM.
- Procesador Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz.
- 40 Gb de disco de almacenamiento.

### 4.4.6.3 Repositorio

Es la base para empezar el despliegue, pues en este lugar se encuentra todo el código de la aplicación y al estar en el gestor de repositorios de GitLab, permite el uso con el Runner.



Name	Last commit	Last update
.vscode	save all registers	
activitiesTimeLineByCourse	order solved	1 month ago
frontend	solved grafos and grap stadistics	3 weeks ago
graphsByDay	solved grafos and grap stadistics	3 weeks ago
graphsBySession	service graphs by session	1 month ago
node_modules	login impelendted	
server	unique email and similarity	1 week ago
similarities	new changes	1 week ago
stadisticTotalByStudent	new config frint and stadistics not graded	1 month ago
stadisticsByStudent	new changes	1 week ago
utils	new changes on services controller	1 month ago
.gitignore	new config to server on selene	6 months ago
.gitlab-ci.yml	deploy similaruty	1 week ago
config.js	new changes on services controller	1 month ago
ecosystem.config.js	new service graphs by day	1 month ago
package-lock.json	script to copy imgs	4 months ago

Figura 4.20 Repositorio en GitLab.

#### 4.4.6.4 GitLab Runner

Es una herramienta software que trabaja con GitLab, que permite interconectar el shell o consola de un servidor o maquina virtual y realizar operaciones. Para poner en marcha la aplicación de SELENEVIS, es necesario instalar el software en el servidor al que se desea conectar [81].

## Runners

Runners are processes that pick up and execute CI/CD jobs for GitLab. [How do I configure runners?](#)

Register as many runners as you want. You can register runners as separate users, on separate servers, and on your local machine. Runners are either:

- **active** - Available to run jobs.
- **paused** - Not available to run jobs.

### Specific runners

These runners are specific to this project.

#### Set up a specific runner automatically

Register a runner on a Kubernetes cluster. [Learn more.](#)

1. Click the button below.
2. Select an existing Kubernetes cluster or create a new one.
3. From the Kubernetes cluster details view, applications list, install GitLab Runner.

[Install GitLab Runner on Kubernetes](#)

#### Set up a specific runner manually

1. [Install GitLab Runner and ensure it's running.](#)
2. Register the runner with this URL:

<https://gitlab.com/>

And this registration token:

5qoc-dK4UzG-QWzsE64D

[Reset registration token](#)

[Show Runner installation instructions](#)

### Shared runners

These runners are shared across this GitLab instance.

[Shared Runners on GitLab.com](#) run in **autoscale mode** and are powered by Google Cloud Platform. Autoscaling means reduced wait times to spin up builds, and isolated VMs for each project, thus maximizing security.

They're free to use for public open source projects and limited to 2000 CI minutes per month per group for private projects. Read about all [GitLab.com plans](#).

#### Enable shared runners for this project



#### Available shared runners: 15

- **pVR9XBDq**  
gitlab-docker-shared-runners-manager-04 #2072991  
[gitlab-org-docker](#)
- **d5ae8d25**  
gitlab-shared-runners-manager-5.gitlab.com #380989  
[gitlab-org](#)

Figura 4.21 GitLab Runner.

#### 4.4.6.5 Servidor HTTP

El servidor web es un software o aplicación que establece los protocolos para la transmisión de información. Es una parte fundamental del despliegue porque genera las conexiones bidireccionales con el cliente, es decir, es el encargado de hacer visible la aplicación al cliente. Para servir la aplicación se optó por usar “Nginx” que es un servidor web de código abierto, usado también como proxy inverso para servir las

interfaces web y la lógica de backend en una misma aplicación, incluso se usa como balanceador de carga en algunos casos.

#### 4.4.6.6 Dominio SELENEVIS

La aplicación se encuentra en el dominio “selene.unicauca.edu.co” con puerto 8000, la figura 4.22 muestra la vista de login.

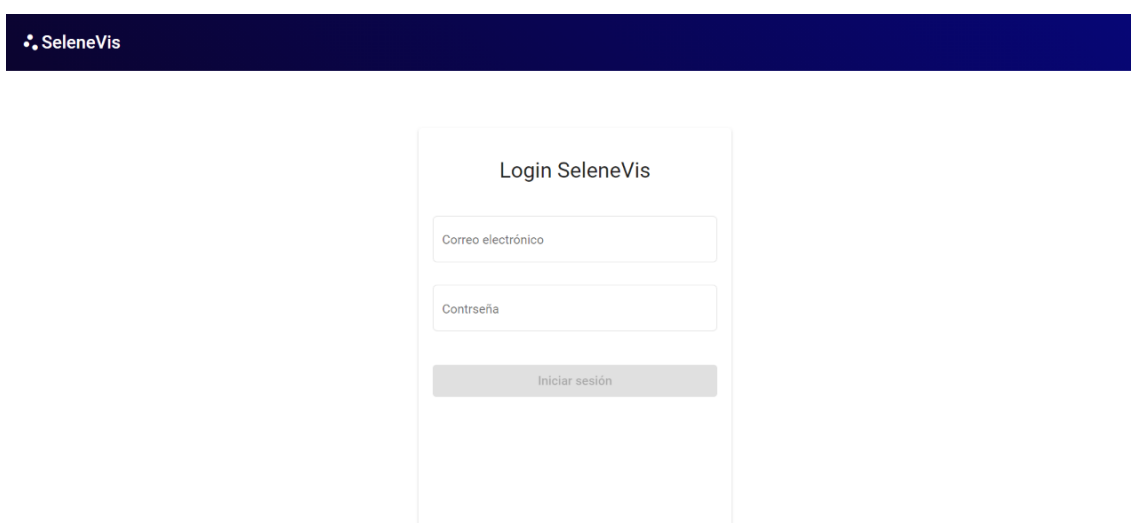


Figura 4.22 Login SeleneVis.

## 4.5 Caso de estudio

En el caso de estudio se ha realizado una descripción del curso, las estrategias propuestas para el mismo y la disposición de las actividades de evaluación de los estudiantes con sus respectivos porcentajes de calificación. Por otra parte, a partir de esta información se obtendrán los resultados tomados a partir del uso de la aplicación Web SeleneVis.



### 4.5.1 Descripción del curso

Para este caso de estudio se ha tomado de referencia el curso “Introducción al emprendimiento con Lean Startup”, curso habilitado el 24 de Agosto de 2020 y con fecha de finalización el día 21 de Diciembre del 2020.

El curso tiene como propósito fundamental, introducir a los estudiantes en el conocimiento y uso de una de las metodologías para el desarrollo de emprendimientos que mayor éxito ha tenido en los últimos tiempos: el Lean Startup.

El curso al ser diseñado desde una perspectiva no presencial, se basa en el desarrollo de actividades de aprendizaje semanales soportadas en las tecnologías de internet, donde el estudiante tendrá acceso a diversos recursos didácticos construidos y seleccionados por el profesor, además de la participación en foros de discusión, la elaboración de mapas conceptuales, uso de simuladores en línea, la presentación de pruebas en línea y el desarrollo de talleres no presenciales.

Durante la primera semana de clase se permite la interacción con los servicios de la plataforma tecnológica que soporta el curso, para que los estudiantes se familiaricen con las funciones y actividades que van a desarrollar durante el curso. Posteriormente, semana a semana, los estudiantes tendrán acceso a un conjunto de recursos digitales que deberán consultar previo al desarrollo de actividades de aprendizaje, algunas de ellas pueden ser calificables y tenerse en cuenta en la obtención de la evaluación del estudiante a lo largo del curso.

El curso fue dispuesto con dos unidades temáticas que conforman un total de 5 módulos que se habilitan semana a semana, de tal forma que en total se desarrollan 5 exámenes correspondientes a cada uno de ellos. Por otra parte, se tiene un examen final que evalúa el total de unidades del curso.

Los exámenes se habilitan 14 días luego de haberse dispuesto el material de cada una de las unidades. Además los estudiantes tienen la posibilidad de presentar un supletorio dos días después del examen oficial. Cada examen es programado a una hora y fecha específica y tienen una duración de 60 minutos.

El examen final dispuesto en el curso, no es el mismo para todos los estudiantes,

dado que se ha asignado un banco de preguntas de las cuales, se asignan de forma aleatoria para formar un examen diferente a cada uno de ellos.

Algunos de los exámenes del curso se componen de dos partes, las preguntas de selección múltiple y actividades que las califica el docente encargado y asigna una respectiva nota. Para el análisis de este trabajo solo se tienen en cuenta las notas de las preguntas de selección múltiple, las cuales las genera la plataforma automáticamente.

# Capítulo 5

## Resultados

### 5.1 SeleneVis

SeleneVis es una aplicación Web que trabaja con los registros de datos de la Plataforma Selene Unicauca, una plataforma para la disposición de cursos en línea y masivos de la Universidad del Cauca. SeleneVis presenta un mecanismo de visualización de datos para la identificación y seguimiento gráfico de los comportamientos con sospecha de fraude de los estudiantes. Está pensada para brindar apoyo a los docentes encargados de los diferentes cursos, para que puedan brindar la asesoría y el acompañamiento a los estudiantes, especialmente a los que presentan comportamientos denominados sospechosos. La aplicación despliega una interfaz del mecanismo de visualización propuesto, y se compone de 4 vistas principales: “Login”, “Sospechosos”, “Estadísticas”, “Comportamientos” y una vista de administración de docentes y/o administradores de la aplicación. A continuación, se describe cada una de las vistas:

#### 5.1.1 Vista “Login”

Al ingresar a la aplicación web la primera vista en mostrarse es la vista de Inicio de sesión (login), en la cual, se requiere del correo electrónico y la contraseña del usuario para poder ingresar a las funcionalidades de la misma. Los usuarios solo

pueden ser registrados a través del módulo de administración descrito en la sección 5.1.5. La figura 5.1 muestra la interfaz web de esta vista.

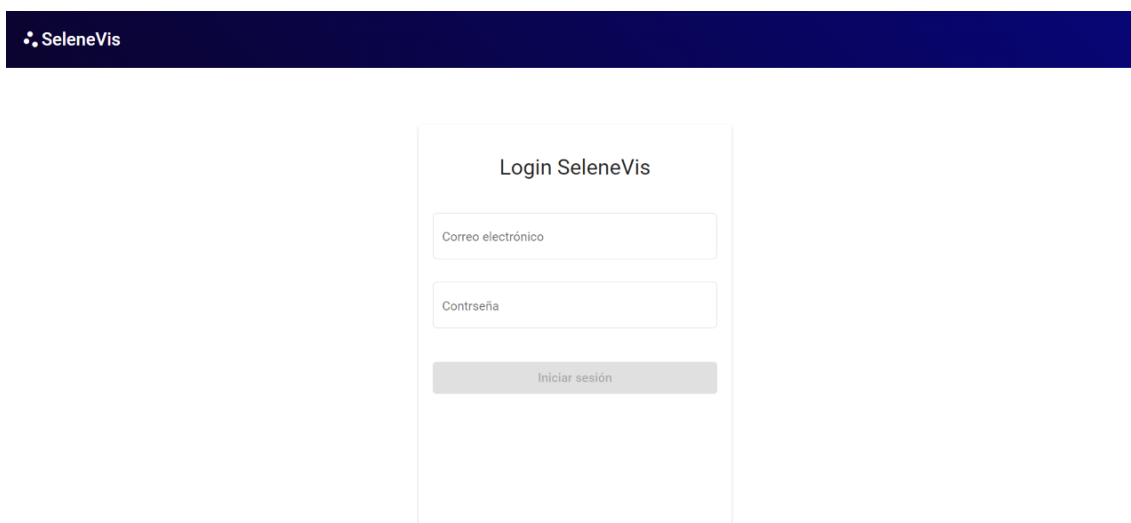


Figura 5.1 Vista de login .

### 5.1.2 Vista “Sospechosos”

El objetivo de esta vista (figura 5.2), es proporcionar al docente una representación gráfica mediante dendrogramas<sup>1</sup> en los cuales, se representa la similitud entre grupos de estudiantes de acuerdo a las respuestas y los tiempos de envío de los cuestionarios, que tuvieron durante el curso o examen que se selecciona en la lista de opciones.

<sup>1</sup>Representación gráfica de datos en forma de árbol, se divide en subcategorías hasta llegar a la información final.

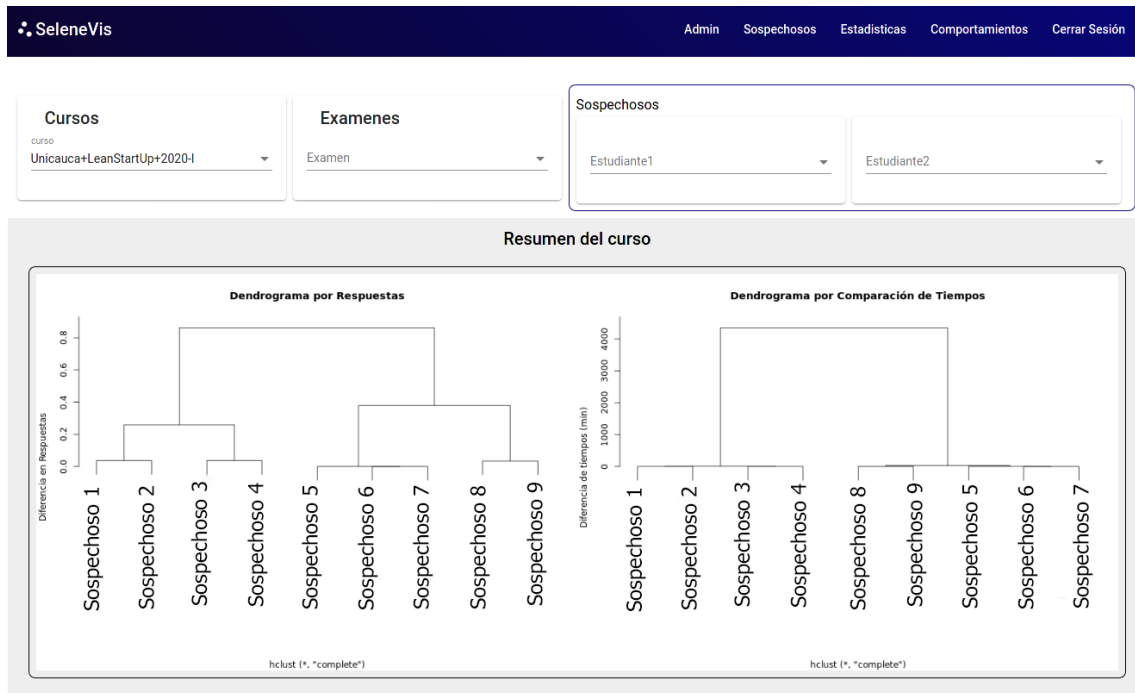


Figura 5.2 Vista sospechosos. Resumen general.

En la parte inferior izquierda se encuentra un dendrograma que representa la cercanía de cada estudiante respecto a las respuestas enviadas. La aplicación agrupa a los estudiante más cercanos. Por ejemplo los Sospechosos 5, 6 y 7 han sido agrupados porque tiene una cercanía de 0. Es decir, tienen un 100% de coincidencia en sus respuestas, tanto correctas como incorrectas. De la misma manera, en la parte inferior derecha se muestra un dendrograma, pero realizado, teniendo en cuenta la cercanía de los estudiantes en el tiempo de envío de sus respuestas. Nuevamente, los sospechosos 5, 6 y 7 han sido agrupados debido a que enviaron sus exámenes muy cerca en el tiempo. La detección de estudiantes sospechosos de fraude y la generación de las gráficas de dendrogramas fue tomada del trabajo previo presentado en [74, 3].

Otra característica de la pestaña de “Sospechosos” es la comparación de dos estudiantes. Esta vista permite obtener directamente la similitud entre estudiantes de acuerdo a su comportamiento en el momento de presentar un examen en la plataforma. La figura [5.3] muestra que al seleccionar el estudiante sospechoso 5 y el sospechoso 6, es posible observar el resultado de su comparación, aquí se evidencia la fecha de

presentación del examen, el número de coincidencias en respuestas y la diferencia de tiempo de envío del examen.

The screenshot shows a web interface for comparing student exam results. It features three main sections: 'Cursos' (Courses) with a dropdown menu showing 'Unicauca+LeanStartUp+2020-I', 'Exámenes' (Exams) with a dropdown menu showing 'Modulo\_2\_Unidad\_Tematica\_2\_Modelo...', and 'Sospechosos' (Suspects) with two dropdown menus for 'Estudiante1' (Student 1) and 'Estudiante2' (Student 2), both set to 'Sospechoso 5' and 'Sospechoso 6' respectively. Below these is a purple 'Consultar' (Consult) button. Underneath the button is a table titled 'Comparación estudiantes' (Student Comparison) with the following data:

Nombre	Fecha de Presentación	# Preguntas del Examen	Coincidencias	Diferencia de Tiempo	Procentaje similitud
Sospechoso 5	10/29/2020, 4:57:49 PM	9	9	0H:0M:3s	100%
Sospechoso 6	10/29/2020, 4:57:46 PM				

Figura 5.3 Similitud estudiantes.

### 5.1.3 Vista “Estadísticas”

Esta vista permite ver las estadísticas generales del curso de acuerdo a las interacciones totales que los estudiantes han tenido dentro del mismo. Las interacciones que se han tenido en cuenta son: interacción con los contenidos, vídeos, foros y exámenes. También se ha tenido en cuenta el número de sesiones que el estudiante presenta en todo el curso. En la figura 5.4 se muestra un pantallazo de dicha pestaña. En la parte izquierda, se muestra un diagrama de barras con la cantidad de vídeos, contenidos, foros, exámenes y sesiones que se han presentado en todo el curso. En la parte derecha, se muestra una tabla con el listado de estudiantes que pertenecen al curso, junto con sus interacciones particulares. Esta tabla puede ser ordenada alfabéticamente por nombre de estudiante o puede ser ordenada de mayor o menor por cada una de las interacciones mencionada anteriormente.

Adicionalmente en la parte inferior, se observa una gráfica con una línea de tiempo, en la cual, se muestran el total de interacciones durante el curso. Esta gráfica presenta la cantidad de interacciones de los estudiantes día a día. Cabe destacar, que esto indica cuál es el comportamiento que tienen los estudiantes durante el curso. Como se puede observar en la parte inferior de la figura 5.4. En el ejemplo de la figura 5.4, en el historial de interacciones, se puede observar con círculos de color

verde los días en que hubo exámenes y se aprecia picos de interacciones un día antes o el mismo día del examen.

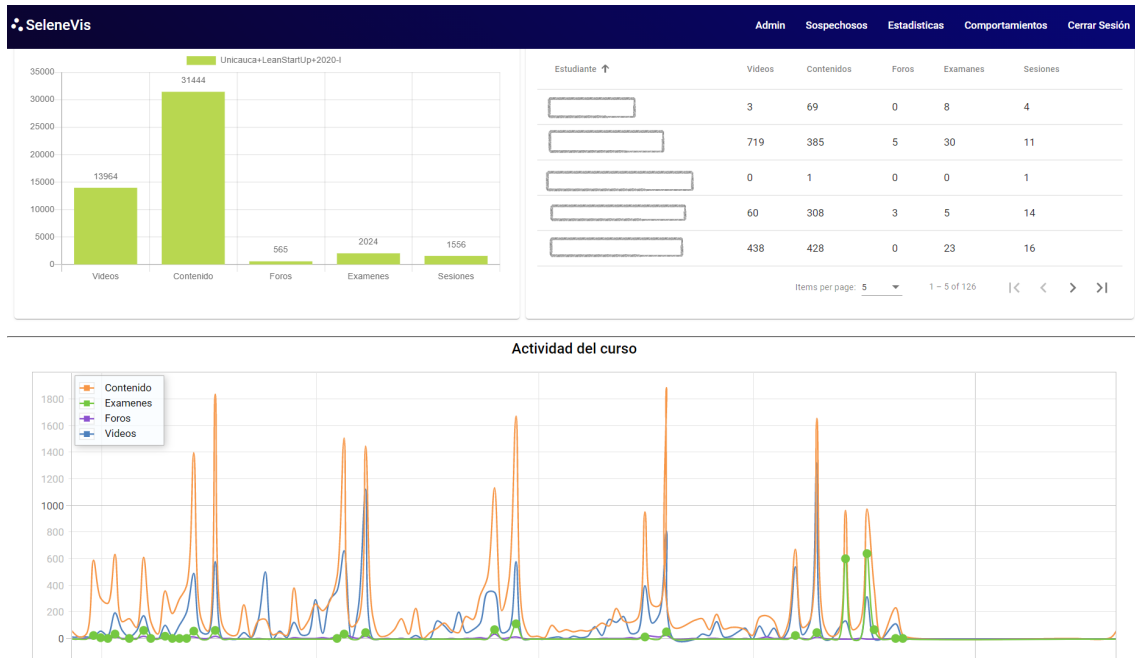


Figura 5.4 Vista estadísticas.

También, dentro de esta vista es posible consultar las estadísticas de un estudiante en particular. El resultado de la consulta se muestra en un diagrama de barras. Solo basta seleccionar un estudiante en la lista y se puede apreciar el resultado en pantalla. La figura 5.5 muestra el resultado de consultar un estudiante.

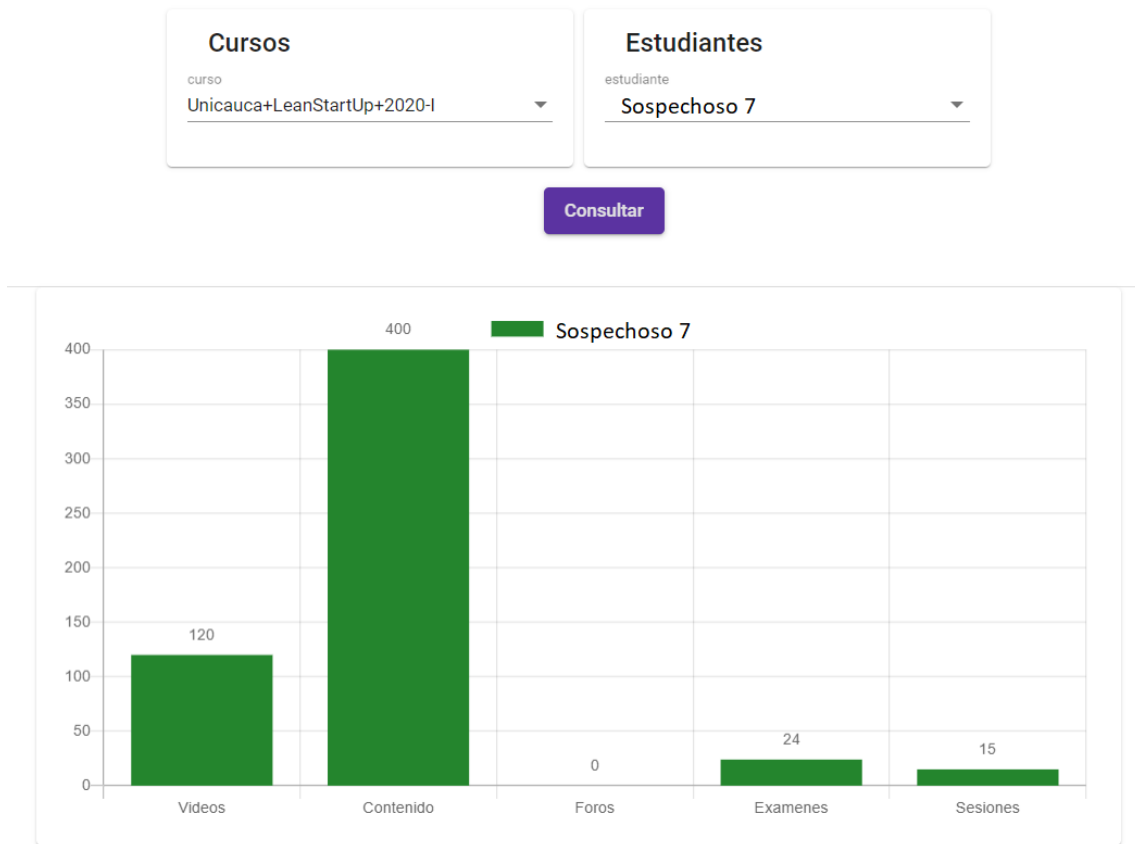


Figura 5.5 Estadísticas por estudiante.

#### 5.1.4 Vista “Comportamientos”

En esta vista es posible consultar y comparar los comportamientos de los estudiantes en el curso. La figura 5.6 muestra los dos tipos de gráficas que se encuentran en esta vista. En la parte superior, están las estadísticas del estudiante consultado y en la parte inferior, se muestra una línea de tiempo orientada a grafos, donde, se observa la relación entre los diferentes nodos (actividades), la cual es un resumen de las actividades que el estudiante realizó dentro del curso, y describe una secuencia ordenada de los pasos que hizo el estudiante, bien sea por día o por sesión.

Una vez consultado el estudiante, encima de las gráficas aparece la opción para escoger la forma en como se desea mostrar los resultados. Es decir, si se escoge la



categoría de filtro por “día”, aparecerá la lista con opciones de los días en que el estudiante interactuó con el curso; y en caso de escoger “sesion”, aparece la lista de sesiones que tiene el estudiante. Una vez escogida la opción se despliega la información estadística y su comportamiento representado por grafos. El gráfico de representación de comportamientos por grafos está basado en el trabajo en desarrollo del estudiante de maestría de la Universidad del Cauca, Luis Alejandro Ordoñez. Él propuso visualizar los comportamientos de los estudiantes mediante grafos, nosotros llevamos su propuesta a la implementación en un desarrollo web. Para ello, se ha dotado la aplicación para realizar las gráficas de manera automática y en tiempo real, algo que no se encuentra registrado en la literatura.



Figura 5.6 Vista comportamientos .

### 5.1.5 Vista “Admin”

La vista “Admin” (administrador), presenta una interfaz que permite listar usuarios registrados y la gestión de nuevos usuarios. Para ello se tiene un formulario de registro y asignación de roles. Estos roles son de docente y administrador. Para las respectivas opciones de registro, edición y eliminación de usuarios, se tiene un formulario que lista los usuarios registrados y además añade los botones para la gestión. Este procedimiento es obligatorio para el posterior inicio de sesión de ingreso a la aplicación. El administrador es quien crea las cuentas de los docentes y asigna los cursos según se encuentren asignados en la plataforma Selene. De esta manera los docentes tienen acceso sólo a los cursos en los cuales estén a cargo.

Nombre	Correo	Cursos	Perfil	Operación
Docente 1	Correo@mail.com	10 cursos	Administrador	
Docente 2	Correo@mail.com	1 cursos	Docente	
Docente 3	Correo@mail.com	9 cursos	Administrador	
Docente 4	Correo@mail.com	1 cursos	Docente	
Docente 5	Correo@mail.com	1 cursos	Docente	
Docente 6	Correo@mail.com	2 cursos	Docente	
Docente 7	Correo@mail.com	1 cursos	Administrador	
Docente 8	Correo@mail.com	3 cursos	Administrador	

Figura 5.7 Vista de Administrador.

## 5.2 Resultados sobre el Caso de estudio

A continuación se muestran los resultados obtenidos del uso del mecanismo de visualización sobre el curso “Unicauca+LeanStartUp+2020-I” descrito en la sección [4.5.1](#).

Inicialmente se analizan las estadísticas generales del curso en la vista de “Estadísticas”.

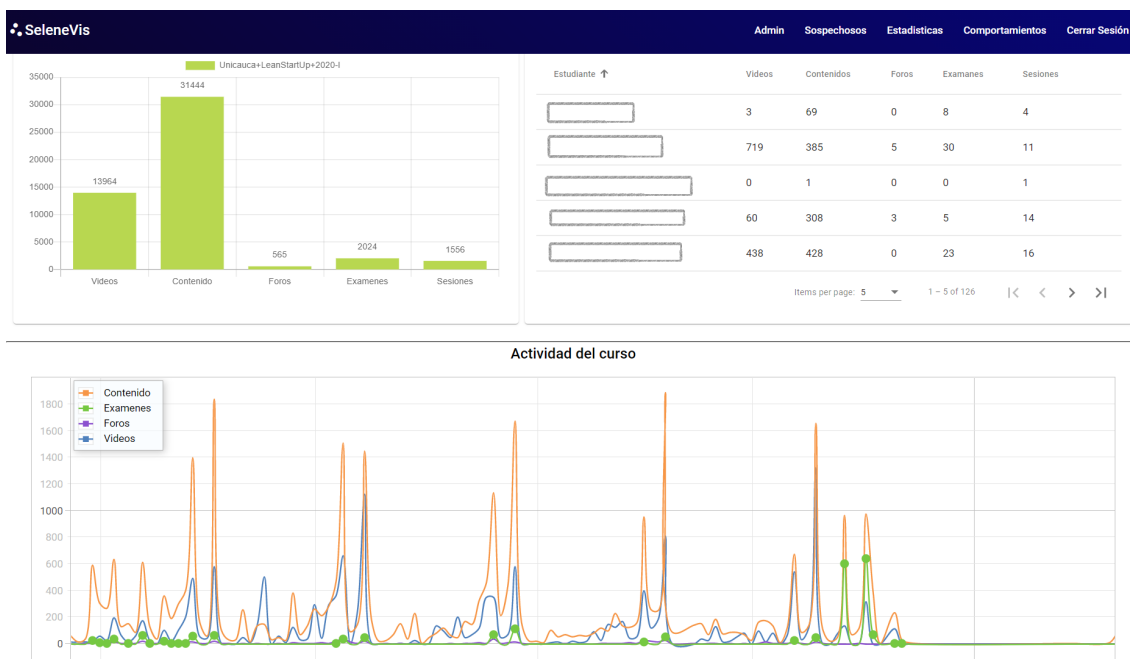


Figura 5.8 Curso LeanStartUp.

De esta vista general es posible identificar las estadísticas globales de la interacción de los estudiantes con las actividades del curso. En la imagen 5.8 se puede apreciar que la participación del curso está dividida de la siguiente manera; 29,1% (13964 interacciones) de interacciones con videos, 65,5% (31534) corresponde a actividades de navegación, lecturas, textos, imágenes y pdf; 1,2% (565) corresponde a interacciones con foros y 4,2% (2024) con actividades relacionadas a los exámenes.

Desde las estadísticas mostradas en la gráfica de línea de tiempo, es posible evidenciar el comportamiento de los estudiantes con las actividades del curso. Se puede apreciar que los estudiantes manejan picos de actividad y estos se dan sobre las fechas en las que se tiene programado los exámenes. Como los exámenes tienen programado un supletorio luego de dos días, es posible verificar que se tiene dos picos continuos que corresponden a la fecha de examen y a su respectivo supletorio; permitiendo descubrir la tendencia de los estudiantes a realizar las actividades de evaluación en los días de supletorio.

Además, es posible evidenciar un crecimiento de actividades con contenido y videos

previo a las fechas de exámenes y de igual forma ver que lo que más realizan los estudiantes es revisar el contenido del curso previo a realizar los exámenes, siguiendo un porcentaje menor de interacciones con los vídeos. De igual forma la proporción de interacción con foros es menor.

También se puede observar que en las fechas del examen final programadas para el 14 y 17 de diciembre, las interacciones de los estudiantes con los exámenes alcanzan un pico elevado que es muy cercano al nivel de las interacciones de los estudiantes con los contenidos del curso que representa el mayor número de interacciones. Se intuye que esto sucede porque el examen tiene un mayor número de preguntas, las cuales fueron 19 y fueron programadas para ser respondidas una a una. Es decir por cada pregunta, los estudiantes consultaban los contenidos. Finalmente es importante resaltar que este examen final no se puede analizar desde la vista de sospechosos propuesto en “SELENEVIS”, puesto que se hicieron exámenes diferentes para la evaluación final de cada estudiante, esto mediante una estrategia de Banco de preguntas, donde se asignan preguntas en forma aleatoria a cada estudiante. Se programó así el examen, porque el docente decidió ver que sucedía si las preguntas eran diferentes para cada estudiante. Lo que se esperaba, es que los estudiantes al no poder compartir sus respuestas tuvieran una mayor interacción con contenidos. Sin embargo no fue ese el resultado, sino que hubo una menor interacción con los contenidos. Se creó que esto sucedió, debido a que la mayoría de los estudiantes ya habían alcanzado la nota necesaria para pasar el curso y ya no mostraron un gran interés en pasar el examen.

### 5.2.1 Identificación de sospechosos de fraude en SELENEVIS

Una vez en la vista “Sospechosos” y seleccionado el curso “Unicauca+L. StartUp+2020-I”, la aplicación resalta los estudiantes que se identificaron como sospechosos de fraude a lo largo del curso. Quienes pueden ser observados en el resumen general dispuesto en la figura [5.2](#).

Estos resultados corroboran el trabajo “Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics” trabajo

realizado por el estudiante de Doctorado Daniel Jaramillo Morillo y el Doctor José RuiPérez Valiente [1]. Debido a que al aplicar el algoritmo propuesto en el trabajo previo, se obtienen resultados similares al ser aplicado sobre dos cohortes diferentes del mismo curso. Es decir los cursos tipo MPOC de la Universidad del Cauca tienen un mismo comportamiento en cuanto a deshonestidad académica detectada mediante la comparación de las respuestas de los estudiantes y el tiempo de envío de sus exámenes.

En la figura 5.9 se muestra 4 grupos de estudiantes, donde los integrantes de cada grupo coinciden en respuestas y tiempos de envío de exámenes. Es de aclarar, que los estudiantes que aparecen en el resumen general, se obtienen teniendo en cuenta todas las preguntas del curso. Para proteger la privacidad e integridad de los estudiantes se realiza un proceso de anonimización de los datos.

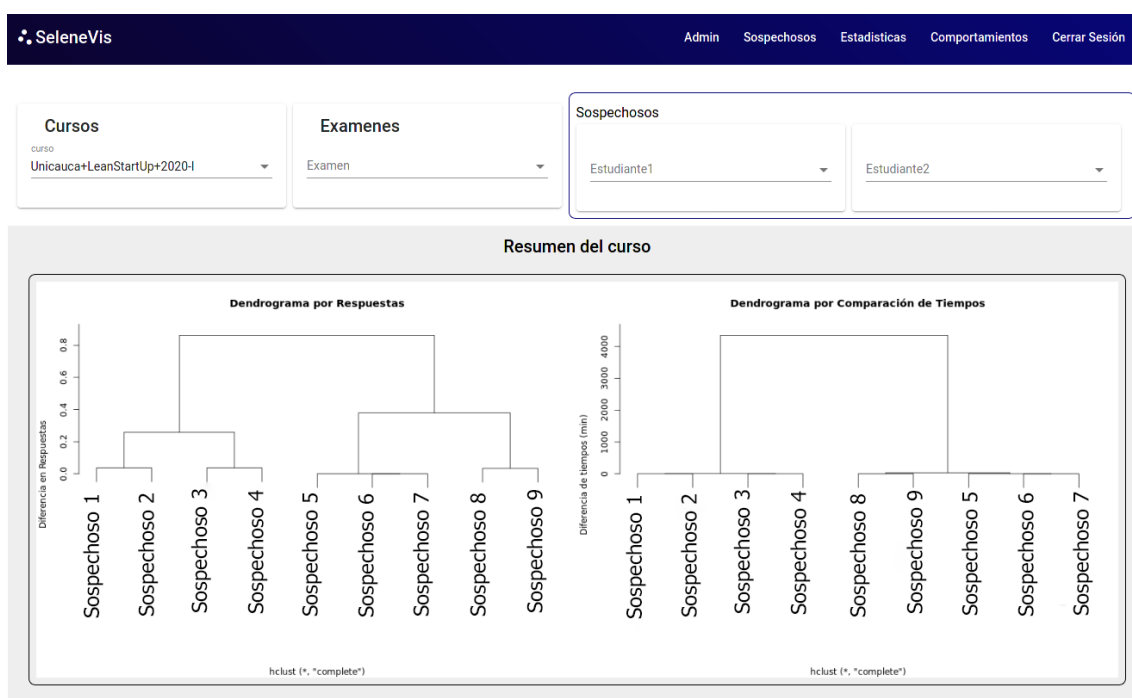


Figura 5.9 Resumen general.

Realizando una exploración en la vista de sospechosos para los diferentes exámenes correspondientes a las respectivas unidades temáticas del curso, se puede identificar nuevos grupos de sospechosos, estos no aparecen en el resumen debido a las consideraciones

del trabajo [1], donde se requiere un porcentaje mayor al 95% de similitud a lo largo de las diferentes actividades de evaluación realizadas en el curso, interceptado con la condición de diferencias de tiempo menores a 5 minutos. Esto siguiendo el procedimiento mostrado en la figura 5.10.

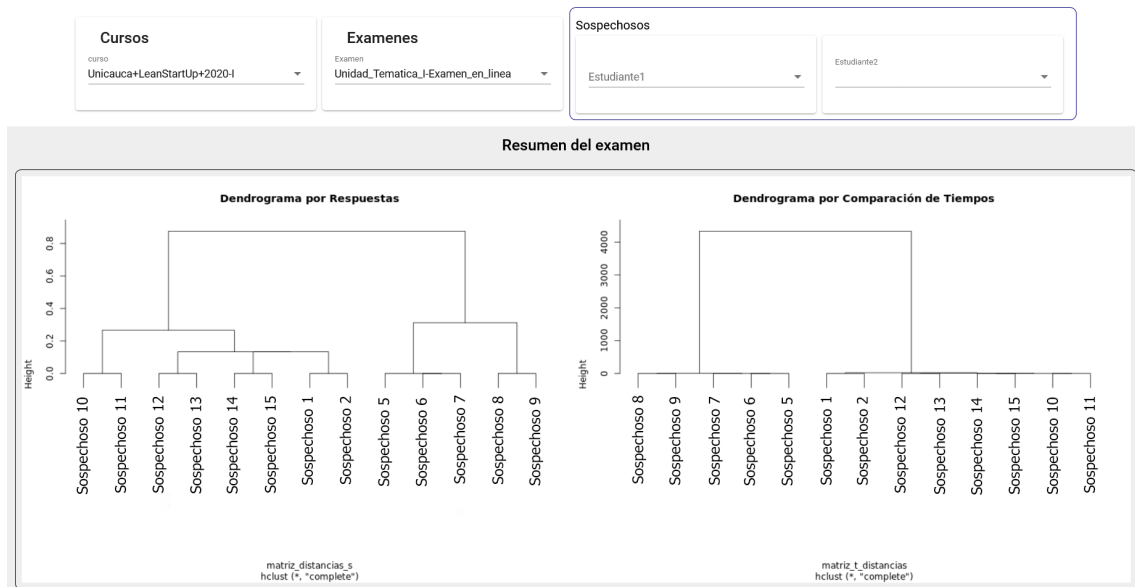


Figura 5.10 Unidad temática I.

Una vez se ha concluido la revisión por cada uno de los módulos en la vista de sospechosos, se puede determinar la lista completa de los estudiantes que fueron identificados por la aplicación como sospechosos, y determinar cuales de los grupos identificados son recurrentes a lo largo de las actividades de evaluación.

Por otra parte, se encontró mediante la vista “sospechosos” el porcentaje de similitud de respuestas entre estudiantes y la distancia de envío en tiempo entre los mismos. Esto siguiendo el procedimiento de la figura 5.3. Con los datos encontrados anteriormente, se realiza una tabla de forma manual y se logra identificar el conjunto completo de sospechosos y la información respectiva de similitud y tiempo, las cuales se condensan en las figura 5.11. En esta figura sólo se muestran los resultados de los estudiantes que la aplicación detectó examen tras examen.

		E1	E2	E3	E4	E5	
1	S1	100			100	100	E1 Unidad Tematica I
	S2						E2 Modulo 1 Unidad Tematica 2: Metodologia Desarrollo de Clientes
2	S3		100		100		E3 Modulo 2 Unidad Tematica 2: Modelo de Negocio
	S4						E4 Modulo 3 Unidad Tematica 2: Validacion de problemas
3	S5	100	100	100	100	100	E5 Modulo 4 Unidad Tematica 2: Producto minimo viable
	S6						
	S7						
	S8						
	S9						
4	S8	100			100	100	
	S9						
5	S10	100					
	S11						
6	S12	100					
	S13						
7	S14	100					
	S15						
8	S16		100		100		
	S17						
9	S12			100			
	S18						
10	S19					100	
	S20						
11	S21					100	
	S22						

Figura 5.11 Sospechosos con porcentaje de similitud obtenidos de la aplicación.

La figura 5.11 condensa y describe el total de 11 grupos de sospechosos que fueron identificados por la aplicación a lo largo de todos los módulos de evaluación E1, E2, E3, E4, E5. Se identifica también algunos grupos que no aparecen en el resumen general, como por ejemplo los grupos 5 al 11.

Anexo a esto, se observa el total de 22 estudiantes sospechosos a los que se ha denominado s1,s2...s22. Como resultado, la tabla permite diferenciar las relaciones que se establecen entre estudiantes a partir de los diferentes grupos que pueden estar conformados por dos o tres integrantes como en el caso del grupo 2 y 3 respectivamente.

Es claro que en algunos grupos las relaciones de estudiantes se mantienen a lo largo del curso, específicamente en los grupos del 1 al 4 que corresponden a los señalados en el resumen general. Estos grupos son identificados como sospechosos en dos, tres o incluso la totalidad de los módulos de evaluación, lo cual se puede evidenciar en los

campos resaltados en la intersección de la fila de grupos y la columna de exámenes de la figura [5.11](#).

Los campos resaltados también cuentan con un porcentaje de similitud entre respuestas de cada estudiante en cada uno de los exámenes, siendo el 100% un total idéntico de respuestas seleccionadas, que incluyen respuestas correctas como incorrectas.

Es importante aclarar en este momento, que todo lo expuesto en la figura [5.11](#), es obtenido de la interacción con SELENEVIS, y que los resultados ahí condensados solo contemplan los sospechosos identificados por la aplicación con sus consideraciones de similitud de respuestas y tiempos de envío, que corresponden a similitudes mayores a 95% y tiempos de envío menores a 5 minutos. [1](#).

Con el fin de hacer una exploración más profunda en cuanto al comportamiento de los grupos de estudiantes e identificar las causas del por qué en algunos exámenes no son detectados como sospechosos, se realiza una exploración manual de las similitudes y los tiempos de exámenes para todos los módulos y todos los grupos, lo cual se describe en la figura [5.12](#), en donde se muestran las tablas en las que se condensó los resultados del análisis manual.



		E1	E2	E3	E4	E5
1	S1	100	92,86	100	100	100
	S2					
2	S3	93,75	100	100	100	92,3
	S4					
3	S5					
	S6					
	S5	100	100	100	100	100
	S7					
	S6					
4	S8	100	87,5	100	100	100
	S9					
5	S10	100	0	0	69	86
	S11					
6	S12	100	85,7	88,8	57,1	100
	S13					
7	S14	100	28,6	100	100	100
	S15					
8	S16	81,3	100	100	100	84,6
	S17					
9	S12	100	92,9	100	57,1	100
	S18					
10	S19	25	0	100	38,5	100
	S20					
11	S21	23,5	0	22,2	93,3	100
	S22					

		E1	E2	E3	E4	E5
1	S1	0	6	852	18	1
	S2					
2	S3	12	15	30	16	0
	S4					
3	S5	1	12	4	8	6
	S6					
	S5	5	10	8	23	2
	S7					
	S6	4	2	4	15	8
4	S8	3	2	31	0	1
	S9					
5	S10	910	258527	160374 8544	259132	1901
	S11					
6	S12	1	62	16	28	25
	S13					
7	S14	0	1030	1359	1600	1560
	S15					
8	S16	635	9	35	9	9
	S17					
9	S12	61	33	2	2	62
	S18					
10	S19	260762	259488	43	259158	2
	S20					
11	S21	1440080 582	261561	260988	446	2
	S22					

Figura 5.12 a. Porcentaje de similitud b. Diferencia de tiempo (segundos).

En la tabla ubicada a la izquierda de la figura 5.12, se resaltan todos los grupos cuyos estudiantes tienen una similitud del 100%, y en la derecha de la misma se resaltan los grupos con diferencia de tiempo mayores a 5 minutos. El valor en la figura b se encuentra en segundos.

De la figura 5.12, es posible señalar que existen 3 casos a identificar:

Los casos que identifica la aplicación (señalados en la figura 5.12), en la cual se encuentra que solo detecta los estudiantes con 100% de similitud y una diferencia de tiempo de envío menor a 25 segundos.

El segundo caso que se logró contemplar es que las diferencias de tiempos de envío entre los 25 segundos hasta 5 minutos, hace que los estudiantes no sean identificados como sospechosos en [1]. Esto debido a que el algoritmo implementado se redujo

el tiempo de detección de sospechosos a tiempos menores 25 segundos para el caso de estudio de la tesis de doctorado del estudiante Daniel Jaramillo. Estudio que es apoyado por nuestro trabajo. Finalmente en el tercer caso, se encuentra los estudiantes que tuvieron 100% de similitud de respuestas y se descartaron por tiempos de envío mayores a 5 minutos.

		E1	E2	E3	E4	E5
1	S1	4,35	3,35	3,9	4	4,55
	S2	100	2,9 92,86	3,9 100	4 100	4,55 100
2	S3	4,35	3,75	4,45	4	2,9
	S4	93,75	3,75 100	4,45 100	4 100	2,9 92,3
3	S5	3,45	2,7	3,9	2,5	3
	S6	100	2,7 100	3,9 100	2,5 100	3 100
	S7	3,75	2,7	3,9	2,5	3
4	S8	2,8	3,1	3,9	2,5	2,9
	S9	100	87,5 3,45	3,9 100	2,5 100	2,9 100
5	S10	5	4,25	0	3,5	4,1
	S11	100	0 3,75	0 3,9	69 3	86 5
6	S12	4,35	3,75	3,9	3	5
	S13	100	85,7 3,75	88,8 4,45	57,1 3	100 5
7	S14	3,65	2,5	2,8	2,5	3,75
	S15	100	28,6 1,65	100 2,8	100 2,5	100 3,75
8	S16	4,45	4,6	3,35	4	3,35
	S17	81,3	100 4,6	100 3,35	100 4	84,6 2,91
9	S12	4,35	3,75	3,9	3	5
	S18	100	92,9 3,35	100 3,9	57,1 3	100 5
10	S19	4,35	2,1	3,9	3	3,6
	S20	25	0 2,1	100 3,9	38,5 3	100 3,6
11	S21	4,05	1,15	1,1	3,5	2,08
	S22	23,5	0 2,08	22,2 3,35	93,3 3	100 2,08

	Detectados por la aplicación
	No detectados tiempo menor a 5 minutos
	No detectados tiempo mayor a 5 minutos

Figura 5.13 Caracterización sospechosos por nota de examen y similitud.

En la figura 5.13, se ha realizado un proceso de resumen y caracterización que identifica los tres tipos de casos que se han señalado con anterioridad. Para tener mayor claridad a la hora de definir si la similitud no se debe a un resultado ideal en el que los estudiantes coinciden en tener todas las preguntas correctas, se ha optado por anexar en la misma figura las notas que los estudiantes obtuvieron en los diferentes exámenes.

Las notas que se adicionan, junto con el porcentaje de similitud permiten identificar

que los estudiantes en cada grupo respondieron de igual manera los cuestionarios sin que necesariamente esto signifique que obtienen una buena nota, evidenciando que respondieron de la misma manera, tanto las respuestas correctas como las incorrectas.

Por otra parte, los diferentes casos se representan en el gráfico a partir de diferentes colores, tal como señalan en la figura 5.13. Siendo categorizados como Sospechosos detectados por la aplicación, No detectados por tiempo menor a 5 minutos y los No detectados por tiempo mayor a 5 minutos. Para hacer un análisis detallado de los comportamientos ante los diferentes casos identificados, se opta por profundizar el estudio sobre los grupos 3, 4 y 7, en los cuales es posible abarcar las categorías antes mencionadas.

El grupo 3 es seleccionado debido a que todos los estudiantes fueron detectados como sospechosos en todos los exámenes del curso.

De igual manera se selecciona el grupo 4 porque al realizar un análisis manual encontramos que uno de los estudiantes ingresaba con anterioridad a la plataforma y el otro no, lo que nos conduce a un posible caso de parasitismo (free-riding). Además se encuentran 3 casos de exámenes identificados como sospechosos por la aplicación y un caso en el que al presentarse una diferencia entre 25 y 30 segundos en el envío de exámenes, no se identifica de por la aplicación. De igual manera el grupo 7 es seleccionado debido a que es el que presenta un mayor número de casos no identificados dentro de la aplicación porque presentan un tiempo de envío mayor a 5 minutos.

## 5.2.1.1 Estadísticas generales sospechosos

grupo	Sospechoso	Vídeos	Contenido	Foros	Exámenes	Sesiones
1	S1	41	308	7	24	11
	S2	214	281	11	27	22
2	S3	141	228	0	25	9
	S4	317	201	2	25	16
3	S5	54	134	0	25	13
	S6	146	269	2	31	15
	S7	120	400	0	24	15
4	S8	376	397	3	7	19
	S9	88	127	1	6	7
5	S10	164	203	13	24	13
	S11	106	270	10	26	12
6	S12	89	543	7	24	16
	S13	149	695	12	24	14
7	S14	719	385	5	30	11
	S15	479	505	5	31	12
8	S16	113	466	18	25	13
	S17	94	279	13	39	22
9	S12	89	543	7	24	16
	S18	119	732	17	25	65
10	S19	284	288	8	42	12
	S20	70	554	2	27	12
11	S21	58	190	0	26	10
	S22	147	940	15	27	17

Figura 5.14 Comparativo general de estadísticas Grupos 3, 4 y 7.

La figura [5.14](#) condensa las estadísticas generales de todos los grupos sospechosos de fraude, pero se resaltan los grupos a analizar, el grupo 3 que fue identificado por la aplicación en todos los exámenes, el grupo 4 que fue identificado por la aplicación en el resumen general, y el grupo 7, el cual fue identificado manualmente como sospechoso. Dichas estadísticas se representan en la gráfica [5.15](#).

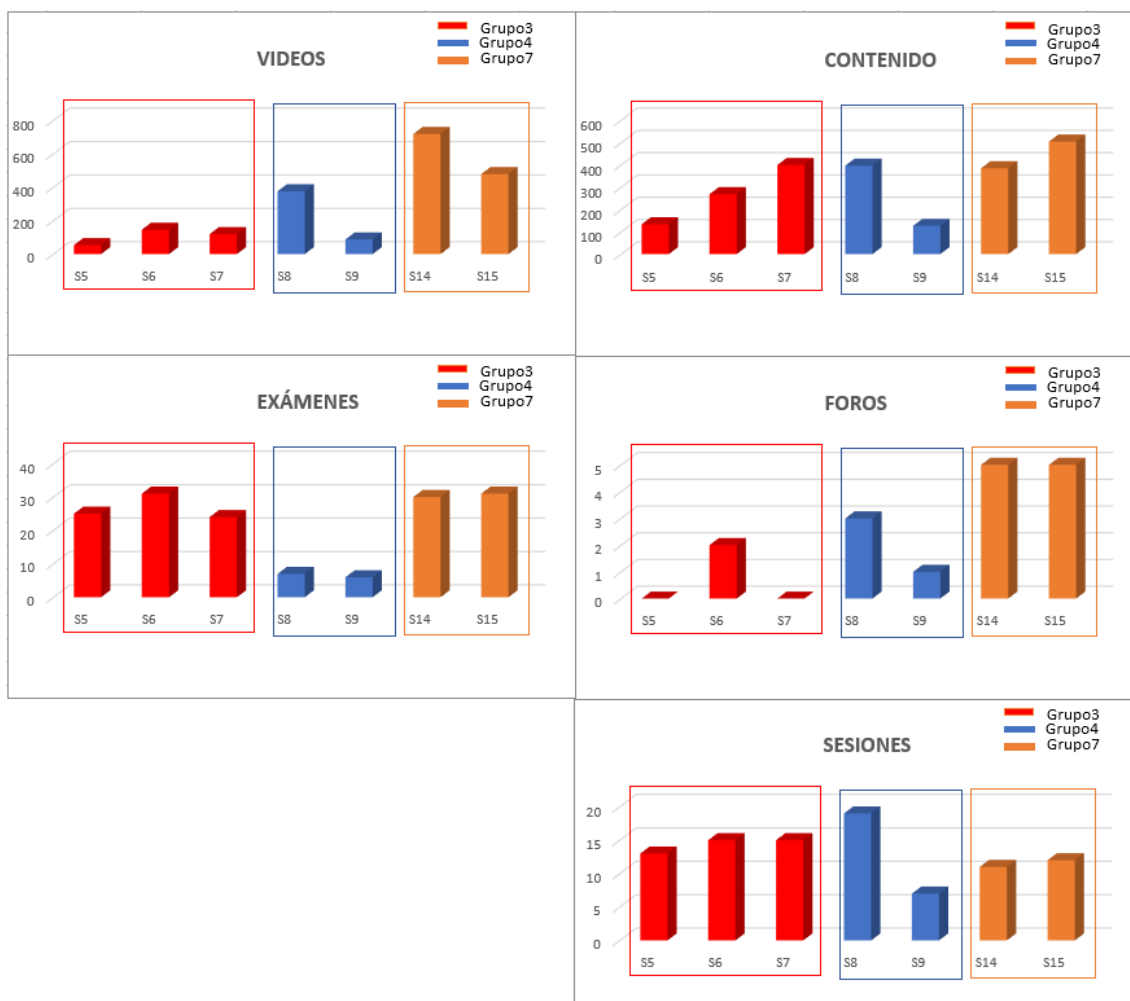


Figura 5.15 Comparativo estadísticas grupo 3, 4 y 7.

En la gráfica [5.15](#) se aprecia que el grupo 7 tiene mayor interacción de contenidos, vídeos y foros, en comparación con el grupo 3 y 4. Sin embargo, el grupo 7 cuenta con un menor número de sesiones, esto significa que este grupo no cierra algunas de las sesiones con las cuales ingresa a la plataforma. Podemos presumir entonces, que los estudiantes de este grupo tienen una mayor interacción con la plataforma, ya sea el día del examen o a lo largo del curso y por eso se mantienen sus sesiones activas y son menores a las del grupo 3 y 4.

Las interacciones de vídeo son muy bajas en el grupo 3 en comparación con los otros grupos 4 y 7, y debido al desempeño en cuanto a notas reflejadas en la gráfica

[5.13](#) y la cantidad de interacciones se presume que el estudiante S5 del grupo 3 está incurriendo en un caso de free-riding, obtiene las mismas notas que sus compañeros pero sus interacciones con los contenidos son mucho menores.

Se observa que en cuanto a exámenes el grupo 4 es el de menor número de interacciones, esto sucede debido a que los estudiantes de este grupo no realizaron los exámenes de bienvenida, los cuales no eran obligatorios dentro del curso. También es posible verificar que el número de interacciones con contenido y vídeos es bajo a lo largo del curso por parte del grupo 4, dado que los estudiantes solo están ingresando en las fechas en las que se tiene programados los exámenes.

También se observa que por lo general en los grupos un estudiante interactúa más en la plataforma y que la diferencia de interacciones entre estudiantes es mucho mayor en el grupo 4. Dando a entender que un estudiante está más comprometido que el otro, este comportamiento en la literatura se lo conoce como free-riding, el cual consiste en que un estudiante desarrolla todo el trabajo y el otro simplemente recibe las respuestas a los exámenes [82](#). En cuanto al grupo 3, es posible evidenciar que la interacción con contenidos de vídeo es baja en comparación con el grupo 7 y 4. Esto quiere decir que se interactúa muy poco con los vídeos del curso, y se puede corroborar comparándolo con el grupo 7 que tiene una interacción constante con el curso, y comparándolo con el grupo 4 que se comprueba que solo interactúa el día del examen.

### 5.2.1.2 Gráficas línea de tiempo

Para entender mejor por qué se presentan esas diferencias en los niveles de interacción entre estudiantes y entre grupos, se realizan las gráficas [5.16](#), [5.17](#) y [5.18](#), en las cuales se representan la cantidad de interacciones con los contenidos, como por ejemplo el material dispuesto en documentos, lecturas, imágenes, navegación y el material dispuesto en vídeos para cada uno de los estudiantes de los diferentes grupos analizados (3, 4 y 7) a lo largo de las diferentes semanas del curso. También en las gráficas se ha señalado con un rectángulo las fechas en las que los estudiantes realizaron el examen, teniendo presente que cada examen tiene su respectivo supletorio luego de 3 días de la fecha asignada.

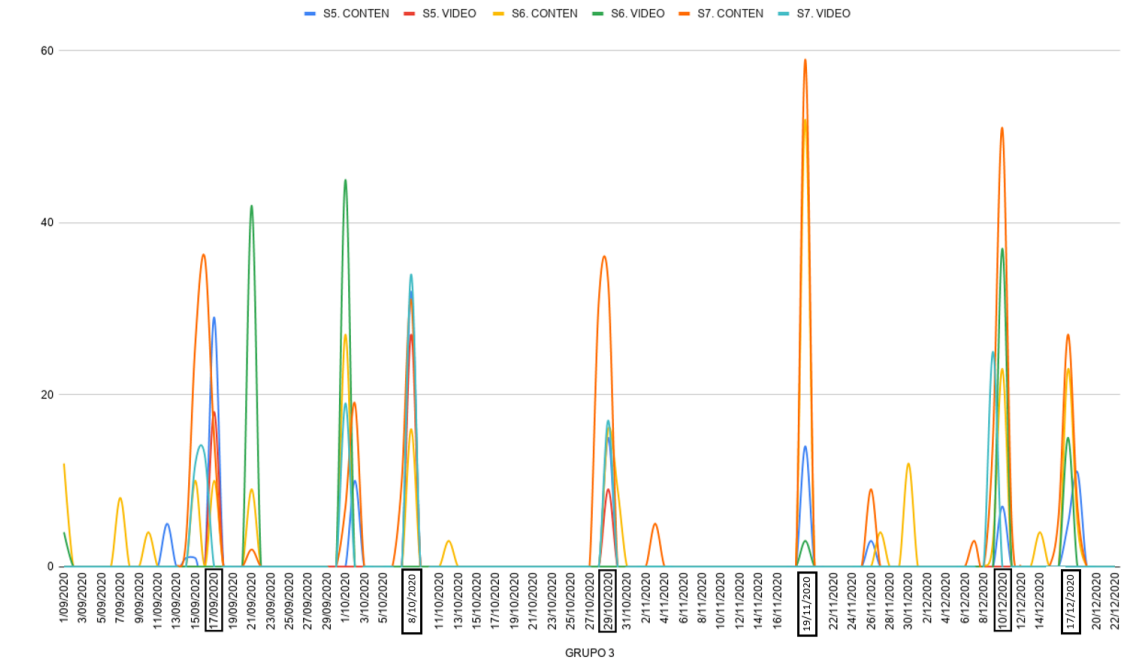


Figura 5.16 Interacción en línea de tiempo grupo 3.

De la gráfica [5.16](#) se puede observar que los estudiantes del grupo 3 tienen la tendencia de estudiar los contenidos días antes del examen a excepción de los exámenes 3 y 4, donde estudiaron el mismo día de la evaluación. Además se observa que las interacciones de los estudiantes son en días muy cercanos, lo que indica que tienden a revisar en espacios similares los contenidos del curso.

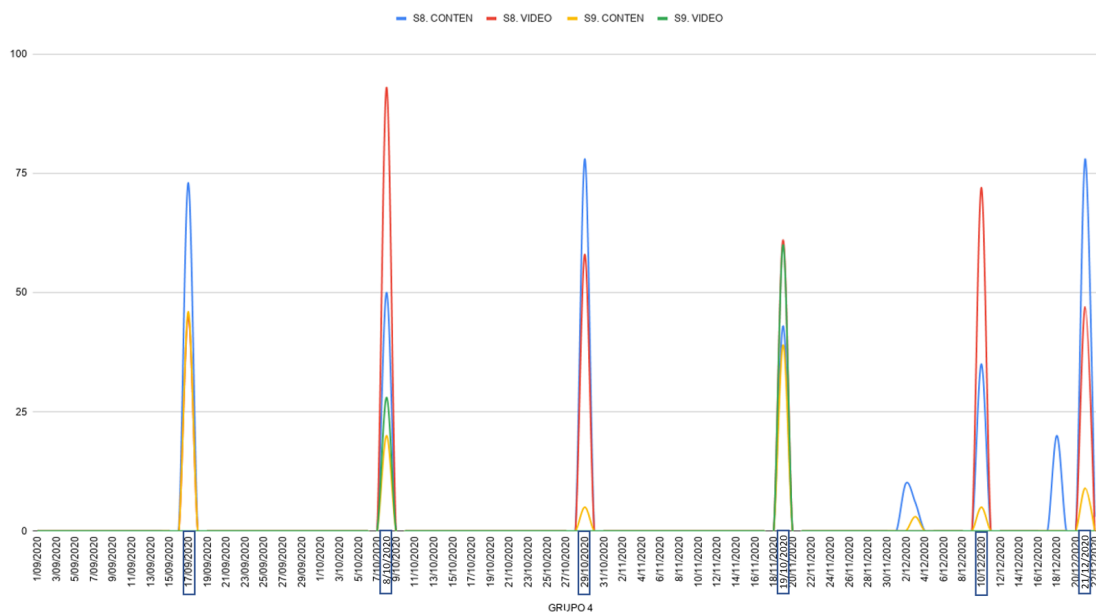


Figura 5.17 Interacción en línea de Tiempo grupo 4.

Podemos observar en la figura [5.17](#), que los estudiantes s8 y s9 solo están ingresando a realizar las actividades del curso el día en el que se tienen programadas las actividades de evaluación. Por otra parte se tiene que al ingresar únicamente el día del examen la nota obtenida para ambos estudiantes es menor a 3, como se ve en la gráfica [5.13](#), que es la nota mínima requerida para aprobar la evaluación. Esto indica que generalmente no se presentan buenas calificaciones cuando se ingresa solo el día del examen, más cuando se evidencia que se ingresa a los contenidos alrededor de la hora programada para el examen, que está comprendida entre las 4:15 y las 5:45 p.m. del día. Para este caso particular del grupo 4, también es posible identificar que el estudiante S9, no interactúa con los vídeos del curso, probablemente porque esta actividad requiere más tiempo para encontrar la información necesaria para responder las preguntas, que revisando una lectura o un archivo PDF, con el fin de responder algo específico del examen una vez se ha iniciado el desarrollo del mismo.

Otro aspecto importante del grupo 4 es que en el examen 3 y el examen 5, los estudiantes revisan el contenido en la hora programada del examen y sobre esa hora realizan toda la interacción representada en la figura [5.17](#). Para obtener esta información se realizó una búsqueda manual en los registros de la base de datos, allí



se encontró que S8 estudia una vez habilitado el examen, mientras que S9 ingresa después de S8 y permanece en la plataforma 16 minutos hasta que ambos envían el formulario con las respuestas.

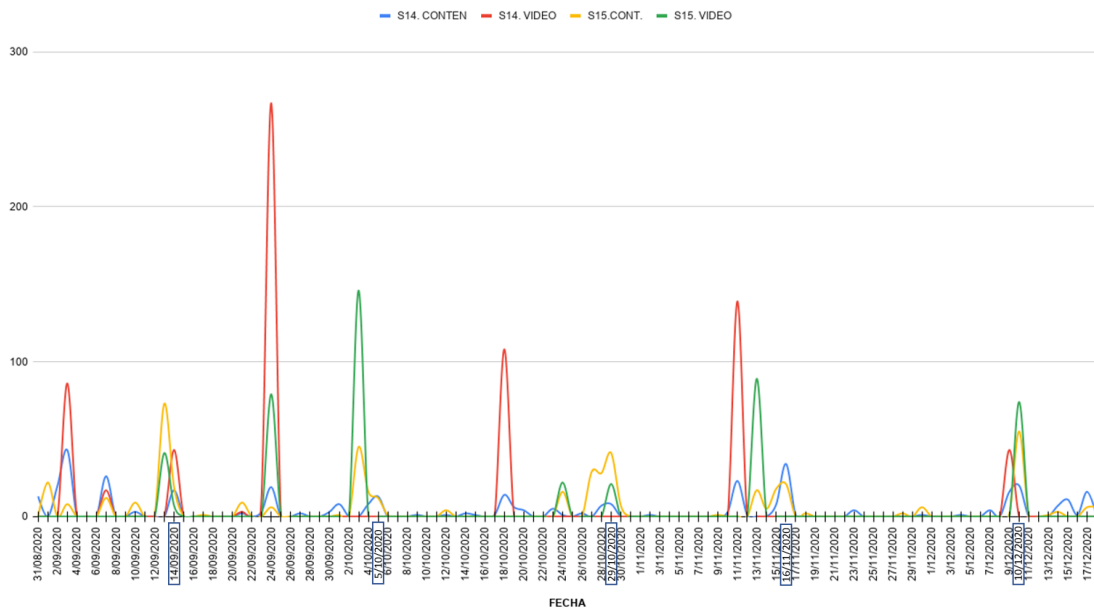


Figura 5.18 Interacción en línea de tiempo grupo 7.

De la figura [5.18](#), se puede corroborar que el grupo 7 tiene mayor volumen de interacciones respecto al grupo 3 y 4, esto visto en la figura [5.15](#), ya que los estudiantes S14 y S15 realizan interacciones con la plataforma de manera recurrente a lo largo de las semanas en las que se han habilitado los diferentes contenidos a ser evaluados. Es posible verificar que los estudiantes ven el material con antelación y presentan un pico de interacciones en el día o los días previos a cada uno de los exámenes programados. Finalmente se observa que el día del examen los estudiantes tienden a tener muy poca interacción con el material y en forma general tienden a revisar más lo relacionado a la categoría de contenidos y en una pequeña proporción los vídeos correspondientes a la temática evaluada.

### 5.2.1.3 Gráfica diferencia de tiempos en envíos

Para poder explorar y profundizar acerca de las diferencias de tiempo de envío de los exámenes por parte de los estudiantes en los diferentes grupos (3, 4, 7). Se ha optado por generar un gráfico que permite evidenciar la diferencia de tiempo de envío del examen entre los estudiantes de cada grupo, además permite evidenciar quién envía primero el examen tal como se aprecia en la figura 5.19.

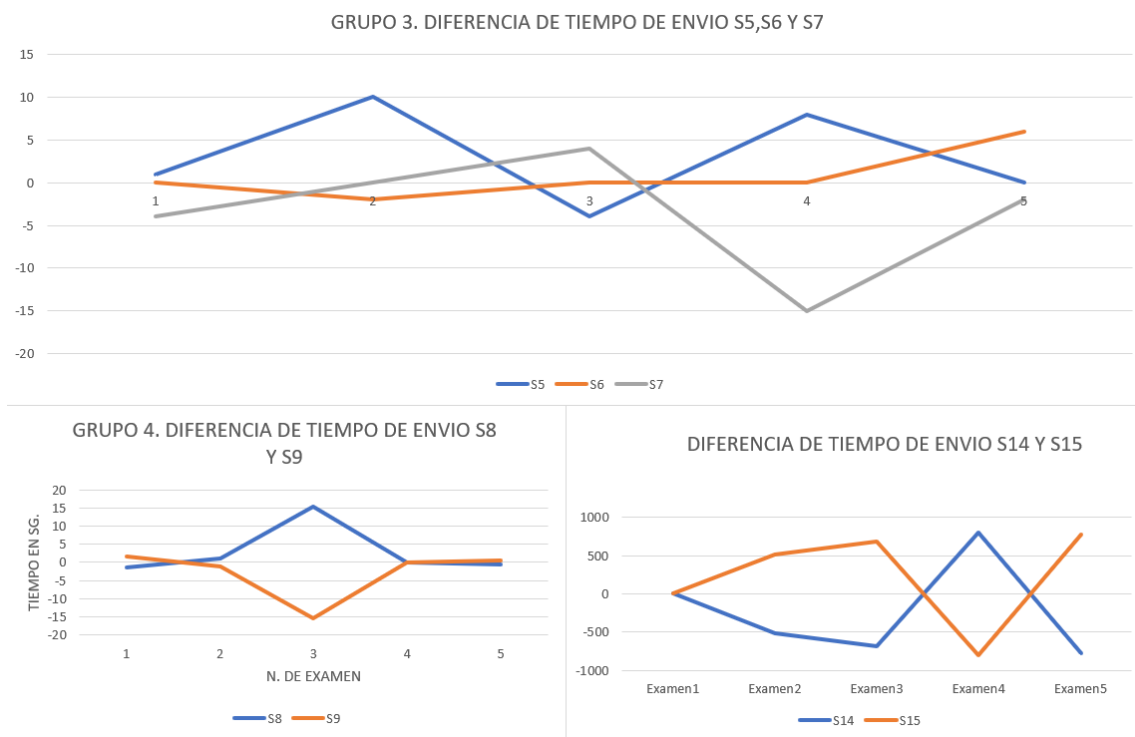


Figura 5.19 Diferencia de tiempos de envío grupos 3, 4 y 7 .

En la figura 5.19 se observan los exámenes donde un estudiante puede depender del otro a la hora de responder el formulario de preguntas.

En este caso encontramos que para tiempos de envío menores a 23 segundos, es indiferente quien envía el examen primero, puesto que se puede comprobar una colaboración en el examen con otro compañero, sin embargo para tiempos mayores a 23 segundos hasta los 5 minutos, es posible evidenciar si se trata de una copia

de respuestas en el examen y quién está respondiendo conforme a las respuestas del estudiante que envía primero. Aquí habría que profundizar en cuál es la razón que determina el tiempo adicional en la plataforma del segundo estudiante en enviar las respuestas respecto al primero.

Para el caso del grupo 3 no se ve un comportamiento muy marcado dado que el tiempo de envío de los exámenes no superan los 30 segundos, es decir que es como si se enviaran al mismo tiempo.

En el caso del grupo 4 se observa que a excepción del examen 3, los estudiantes tienen un comportamiento colaborativo a la hora de presentar el examen, es decir, los estudiantes responden y envían el cuestionario idénticamente y con unos segundos de diferencia, lo que indica que resuelven juntos los exámenes. Sin embargo a pesar de enviarlo juntos, este caso identificó al estudiante S9 como un estudiante que posiblemente este incurriendo en free-riding debido a la baja interacción con el material del curso respecto a su compañero de grupo visto en la figura 5.4, los tiempos de envió, y la hora de ingreso a desarrollar el examen.

En cuanto al grupo 7, es posible verificar de manera notoria la diferencia de tiempo en el que los estudiantes envían las respuestas del examen, es posible evidenciar que se intercalan los envíos y que esos tiempos pueden llegar a estar comprendidos hasta los 30 minutos de diferencia. Lo cual genera incertidumbre de fraude.

#### 5.2.1.4 Esquema de Grafos grupo 3, 4 y 7

Para ofrecer una abstracción del comportamiento de los estudiantes, se ha optado por visualizar mediante la vista de "Comparación" inmersa en la aplicación SeleneVis, la secuencia de interacción con la plataforma de los diferentes grupos de sospechosos. Para ello aprovecharemos de la vista de "comparación" el diseño comparativo mediante el esquema de Grafos que hemos desarrollado y un gráfico de estadísticas complementario. Esto en las fechas en que los sospechosos de cada grupo realizaron los diferentes exámenes.

La exploración de los exámenes se ha hecho teniendo en cuenta los hallazgos de las secciones 5.2.1.1, 5.2.1.2, 5.2.1.3, donde para el grupo 3 se seleccionan el examen

3 y 4, que es donde estudiaron el mismo día de examen y son identificados por la aplicación en el resumen general.

Para el grupo 4 se analiza el examen 3 donde los estudiantes no fueron identificados como sospechosos por la aplicación debido a tiempos mayores a 23 segundos y menores a 5 minutos, y el examen 5 donde los estudiantes si fueron identificados por tiempos menores a 23 segundos por la aplicación. Siendo estos dos casos muy similares en cuanto a tiempos de envío. Además se considera esta selección dado que los estudiantes revisan los contenidos solamente dentro de la hora programada del examen y por otra parte en este grupo 4 se identifica un posible caso de free-riding por parte del estudiante S9.

Para el grupo 7 se escoge el examen 3 y 5, donde no fueron identificados por la aplicación por tener una diferencia de tiempo en el envío del examen mayor a 5 minutos.

### Grupo 3

La figura 5.20 y la figura 5.21 muestran los grafos de comportamiento para los estudiantes en el día del tercer examen, en ellas se observa tanto las estadísticas de interacciones como la secuencia de acciones que cada estudiante realizo en ese día, se observa que los estudiantes S5 y S7 después de entrar a la plataforma van a revisar el contenido y los vídeos del curso antes de presentar el examen, en cambio el estudiante S6 entra a la plataforma revisa los contenido y presenta el examen y no interactuó con los vídeos.

Apoyados en la gráfica 5.16 se puede observar que los estudiantes par este examen solo estudian el mismo día, es decir no se ven interacciones relevantes antes de el día del examen, por tanto se concluye que los estudiantes resuelven el examen solo con lo que estudian ese día.

De lo anterior se puede concluir que el estudiante S6 depende de los otros dos estudiantes, ya que se ve que no estudio en gran cantidad los contenidos y no estudio los vídeos.

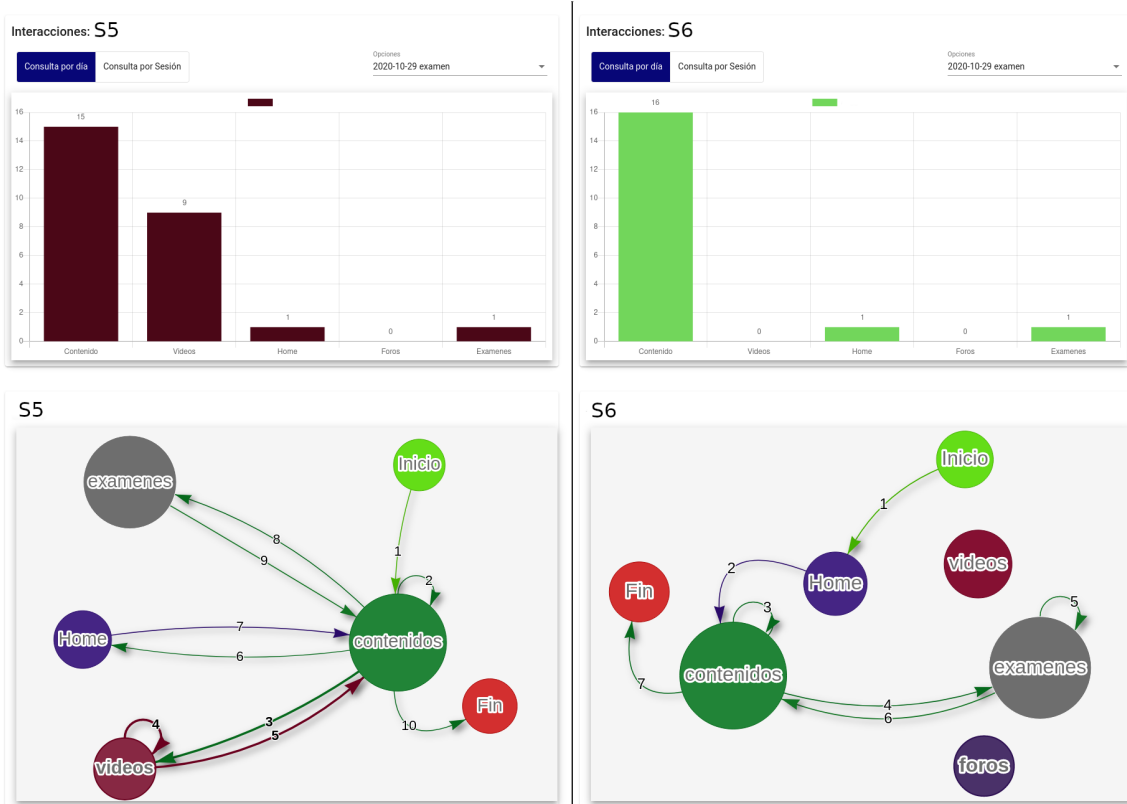


Figura 5.20 Grupo 3 examen 3 S5, S6.



Figura 5.21 Grupo 3 examen 3 S7 .

De las figuras 5.22 y 5.23 las cuales representan las interacciones de los estudiantes en el cuarto examen se puede observar que los estudiantes con más interacciones

son los estudiantes S6 y S7, y el estudiante S5 presenta poca interacción con los contenidos, quienes después de ingresar a la plataforma revisan los contenidos para luego presentar el examen.

de la gráfica 5.16 se aprecia que este grupo no tuvo interacciones notables antes de realizar el examen en cuestión, lo que indica que solo estudiaron el mismo día del examen, por tanto posiblemente el estudiante S5 sea el que depende de los otros dos estudiantes, ya que tiene una cantidad muy baja de interacciones en el día del examen.

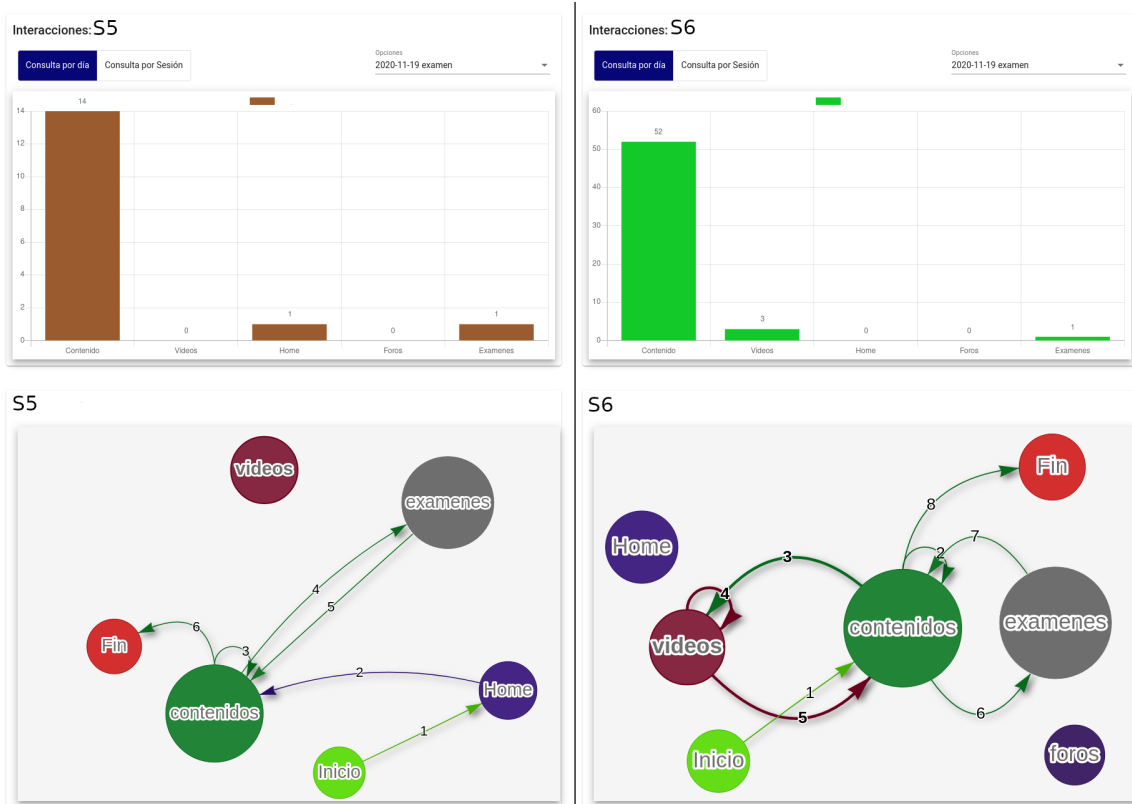


Figura 5.22 Grupo 3 examen 4 S5, S6.

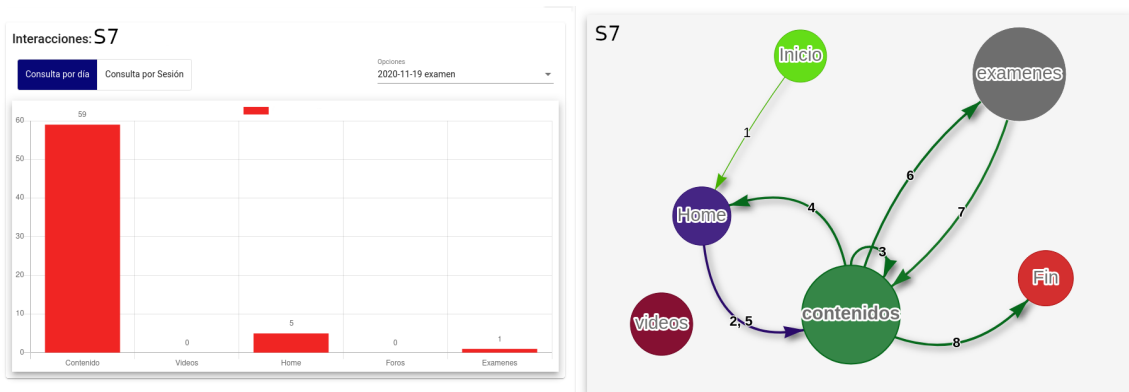


Figura 5.23 Grupo 3 examen 4 S7.

#### Grupo 4

La gráfica [5.24](#) muestra el comportamiento del grupo 4 el día del tercer examen, se observa que el estudiante S8 es el estudiante con más interacciones tanto en vídeos como en contenidos, se ve en la parte de grafos que el estudiante entra a la plataforma revisa los contenidos, luego estudia los vídeos para luego presentar el examen, a diferencia del estudiante S9 quien ingresa a la plataforma, luego tiene muy pocas interacciones con contenidos respecto a su compañero y presenta el examen con el 100% de similitud.

Apoyados en la figura [5.17](#), se puede observar que estos estudiantes solo estudian el día de examen, por tanto en este examen se presenta posiblemente un caso de free-riding por parte del estudiante S9, quien no estudia en la misma proporción del estudiante S8, pero obtiene la misma nota, dando a entender que este estudiante posiblemente copia las respuestas. Y apoyados en la gráfica [5.2.1.3](#) se observa que en este examen quien envía primero las respuestas es el estudiante S8 y posteriormente el estudiante S9, corroborando lo anteriormente mencionado.



Figura 5.24 Grupo 4 examen 3.

En la figura [5.25](#) se observa el comportamiento del grupo 4 en el quinto examen, donde claramente se observa un comportamiento con pocas interacciones por parte del estudiante S9, quien ingresa a la plataforma revisa algunos contenidos y presenta el examen, a diferencia del estudiante S8 quien presenta una gran cantidad de interacción, quien ingresa a la plataforma, revisa contenido, revisa los vídeos en gran cantidad y posteriormente realiza el examen.

Esto da un mayor peso al posible caso de free-riding mencionado en la gráfica [5.24](#) y que se presenta también en este caso.



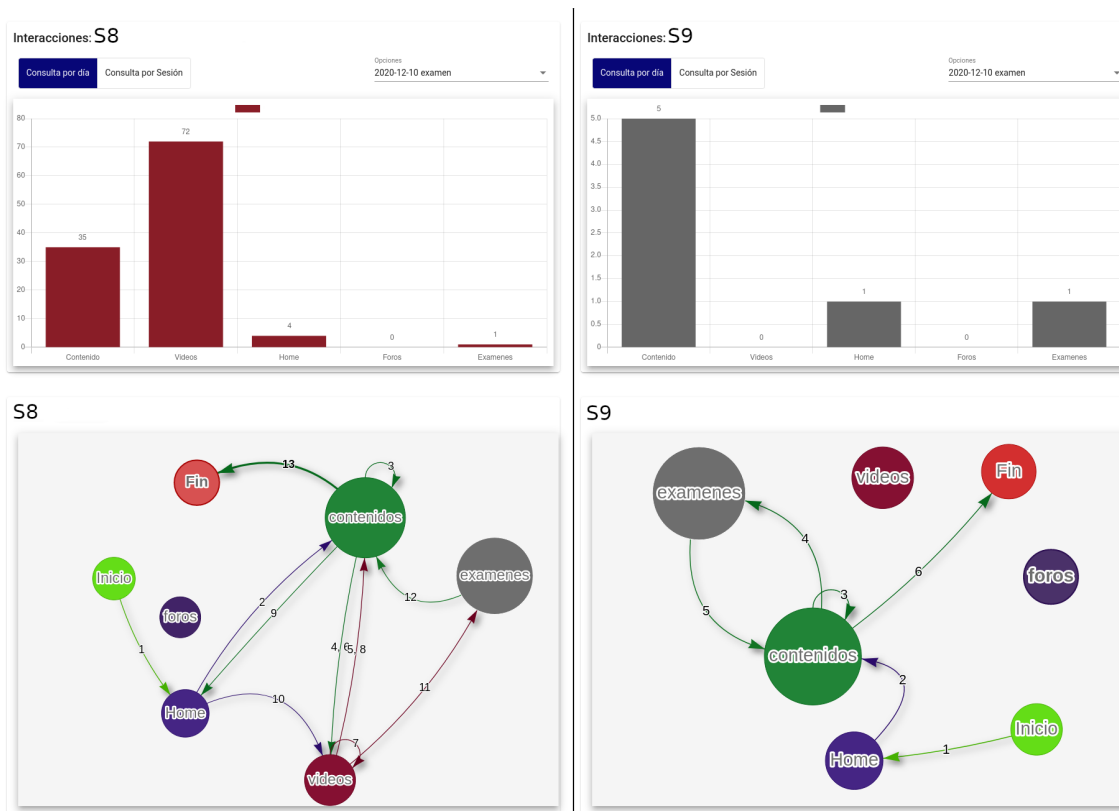


Figura 5.25 Grupo 4 examen 5.

De las gráficas [5.24](#), [5.25](#) se puede concluir que posiblemente dan la evidencia de que el estudiante S9 está copiando las respuestas correctas e incorrectas del estudiante S8, incurriendo en un caso de parasitismo, que consiste en esperar que uno de los estudiantes realice la totalidad de las preguntas del examen y proceder a copiar las respuestas del compañero.

### Grupo 7

La figura [5.26](#) muestra el comportamiento de los estudiantes del grupo 7 en el día del tercer examen, donde se puede observar una baja interactividad por parte del estudiante S14, quien ingresa a la plataforma, revisa una baja cantidad de contenidos y presenta el examen, a diferencia del estudiante S15 quien tiene un alta interactividad en contenidos y en vídeos.

Refiriéndonos a la gráfica [5.18](#) se observa que este grupo estudia días antes de

presentar el examen, por tanto es posible que el grupo presente un comportamiento colaborativo en los exámenes, es decir ambos estudian y resuelven el examen juntos.



Figura 5.26 Grupo 7. Examen 3.

En la figura 5.27 se presenta el comportamiento del grupo 7 en el quinto examen, en el cual se presenta un comportamiento similar al de la gráfica 5.26 y refiriéndonos a la figura 5.18 se puede observar que para este examen el estudiante S14 estuvo estudiando un día antes del examen y el estudiante S15 el mismo día del examen, además presentan a la misma hora el examen y sacan la misma nota. Lo que da a

entender que este grupo trabaja en forma colaborativa para presentar los exámenes.



Figura 5.27 Grupo 7 Examen 5.

### 5.2.1.5 Conclusiones

Finalmente se puede concluir que en los diferentes grupos analizados, por lo general en cada examen existe un estudiante el cual depende del otro, dando como resultados casos de free-riding, como se analizó en el grupo 4. Además, apoyados en la gráfica

**5.13** se concluye que para estos casos de free-riding los resultados en cuanto a las notas no son muy favorables, es decir pierden el examen.

En los casos detectados por la aplicación como sospechosos y en los no identificados, se puede apreciar casos colaborativos entre estudiantes, como lo es en el grupo 7.

También es posible identificar que los estudiantes que interactúan con la plataforma de forma recurrente no tienen mucha interacción con los contenidos ni vídeos en medio del examen, como se ve en las gráficas del grupo 7 para el estudiante S14.

# Capítulo 6

## Conclusiones

La manera en que los MOOC se están incorporando en el ámbito universitario es a través de los MPOC y SPOC, cursos que al ser privados permiten tener un mejor control sobre los estudiantes. Sin embargo, al ser cursos con reconocimiento académico son más susceptibles que los estudiantes cometan fraude en sus exámenes.

Los MOOC al tener reconocimiento académico hace que los docentes necesiten tener un mayor control y seguimiento de sus estudiantes, lo cual no es posible en las plataformas actuales destinadas para MOOC, ya que este tipo de cursos están destinados al auto aprendizaje.

El desarrollo de trabajos como "Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics" desarrollado por el estudiante de Doctorado Daniel Jaramillo-Morillo y el doctor José Ruipérez-Valiente [1], permiten un adelanto en la identificación de estudiantes sospechosos de fraude en el entorno de los SPOCs a partir de métricas de similitud de respuestas y tiempos de envío. Para ello se ha retomado su trabajo y se ha generado una conexión que permite apreciar sus resultados actualizados en el mecanismo de visualización que se ha desarrollado.

Para la implementación del mecanismo de visualización "SELENEVIS", se ha retomado los trabajos de "Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics" [1] y un trabajo de maestría de

la Universidad del Cauca. Para el primer caso se ha realizado una conexión que permite incorporar los resultados en el mecanismo de visualización, haciéndolo más accesible y permitiendo un análisis complementario a partir de datos estadísticos a lo largo del curso. En el segundo caso, se ha retomado una idea de seguimiento de las interacciones de los estudiantes mediante conexiones de nodos. Para lo cual se ha realizado la implementación gráfica de la estrategia en el mecanismo de visualización, a partir de la aplicación Web que permite generar un esquema de grafos interconectados que permiten el seguimiento en línea de manera cronológica de las interacciones de los estudiantes con la plataforma y el material del curso, en las diferentes sesiones o en un día específico a lo largo del curso.

El mecanismo de visualización "SELENEVIS", es una aplicación web que proporciona a los docentes y administradores una herramienta para dar seguimiento a los diferentes cursos que cada uno de ellos tenga asignados. Permite ver el comportamiento de los estudiantes respecto a sus interacciones por sesiones o por días con la plataforma. Agrupa diferentes vistas que permiten diferentes niveles de abstracción. Estas son: la vista de administrador, la vista de sospechosos, la vista de comparación y la vista de estadísticas.

Para la implementación del mecanismo, los datos de los estudiantes se obtienen del archivo tracking.log que ofrece la plataforma SELENE UNICAUCA, instancia de OpenEdx implementada por la universidad del Cauca. Una vez estudiados los campos de este archivo, se retoma el trabajo propuesto por el estudiante de doctorado Daniel Jaramillo Morillo en su tesis de maestría [3]. A partir de esta información se realiza una categorización de estos datos, lo que permite la construcción de los diferentes componentes que finalmente se integran en el mecanismo de visualización. Permite consultar la información en tiempo Real (siendo el tiempo real, consultar los datos actualizados del día inmediatamente anterior al día en el que se hace las respectivas consultas) desde la aplicación web.

Mediante el mecanismo de visualización se puede realizar un seguimiento de las estadísticas generales a nivel del curso o a nivel de cada uno de los estudiantes. También es posible administrar los usuarios docentes y administradores en la plataforma, e identificar estudiantes con sospecha de fraude a partir de las métricas de similitud

de respuestas y similitud de tiempos de envío de los cuestionarios y el seguimiento gráfico de las interacciones en el curso.

Con esta información es posible hacer una revisión específica de los estudiantes detectados como sospechosos a partir de una vista de comparación que muestra las estadísticas de los estudiantes acorde a la categorización de los datos en Contenidos, Vídeos, Foros, Exámenes y Sesiones. Además, esta vista de comparación permite hacer seguimiento de las interacciones de los estudiantes a lo largo de los diferentes días transcurridos del curso o en función de las sesiones. Esto mediante un esquema de grafos conectados que permite seguir el orden en el que interactuó el estudiante.

Mediante la información proporcionada por las diferentes vistas del mecanismo de visualización, se realizó un proceso de identificación y seguimiento de sospechosos, que permitió descubrir, nuevos grupos que no aparecen en el resumen general del trabajo [1], los cuales se denominaron: Estudiantes identificados por la aplicación, Estudiantes no detectados por la aplicación por tiempos de envío entre 25 segundos y 5 minutos, y los estudiantes no identificados por la aplicación por tiempos de envío mayores a 5 minutos. Esto teniendo en cuenta que todos los casos de sospechoso tienen 100% de similitud en respuestas correctas como incorrectas en los diferentes exámenes.

En base a esto fue posible identificar que en el trabajo [1] hay una métrica de tiempo de envío menor a 25 segundos, esta métrica es usada para identificar a los grupos sospechosos de fraude, esto implica que la aplicación no detecta a los sospechosos en algunos exámenes aun cuando tienen 100% de similitud en sus respuestas. Para el caso de estudio que se desarrolló, se encontró que un valor de menos de 31 segundos pudo haber detectado la mayor cantidad de casos sospechosos en los diferentes exámenes. Sin embargo, es necesario profundizar sobre este valor de métrica analizando mayor número de casos de estudio, con el fin de detectar la mayor cantidad de grupos con sospecha de fraude.

Una vez reunida la información proporcionada por el mecanismo de visualización se crearon una serie de gráficas mostradas en [5.2.1.1], [5.2.1.2], [5.2.1.3] y [5.2.1.4]. Estas gráficas permiten diferentes niveles de abstracción, que ayudan a esclarecer el comportamiento de los estudiantes. Esto a partir del gráfico de caracterización de sospechosos, el

gráfico de estadísticas de grupos sospechosos identificados, el gráfico de comportamientos en línea de tiempo, el gráfico de diferencias de tiempo de envío que permite identificar que estudiante envía primero el examen y finalmente el gráfico de esquema de grafos, que visualiza el orden cronológico de las actividades de los estudiantes en la plataforma.

Una vez reunida la información de los diferentes niveles de abstracción se encontró que los estudiantes tienden a tener comportamientos colaborativos, es decir se ponen de acuerdo para desarrollar de manera cercana ciertas actividades del curso, siendo esto más notorio los días previos o el mismo día de la realización del examen y en algunos casos, solo en la hora en la que se habilita el desarrollo del mismo.

Por otra parte, fue posible identificar la existencia de estudiantes que posiblemente están incurriendo en el comportamiento de free-riden. Sin embargo, habría que corroborar este comportamiento haciendo énfasis en estrategias a la hora de asignar las preguntas a los exámenes de cada estudiante, también profundizando en el comportamiento de los estudiantes a través de las horas de ingreso al desarrollo del examen y el respectivo envío del mismo por parte de cada uno de los ellos. Además, es importante ahondar en cuál es la razón que determina que en algunos casos un estudiante identificado como sospechoso envíe el examen mucho tiempo después del envío del cuestionario por parte del primer estudiante, conociendo que en algunos casos pueden iniciar el desarrollo del examen al mismo tiempo y en otros casos no.



# Capítulo 7

## Trabajos Futuros

Se propone como trabajos futuros incorporar de una mejor manera las gráficas [5.2.1.1](#), [5.2.1.2](#), [5.2.1.3](#) y [5.2.1.4](#) en la aplicación “SELENEVIS”, porque en el momento están separadas y las gráficas se complementan entre sí, al estar juntas permitirían analizar mejor la situación de los estudiantes con sospecha de fraude. Una de las recomendaciones es que esta información que proporciona la aplicación se lleve a estas gráficas de manera automática. Uno de los ejemplos más importantes sería tener en un mismo lugar las gráficas de línea de tiempo junto con el esquema de grafos y estadísticas.

Se recomienda en trabajos futuros considerar la nota del examen final y si el estudiante aprobó o no el curso, con el fin de determinar, por ejemplo, si los casos de free-riding ganan o no el examen final, dado que este exámen no es el mismo para todos sino que las preguntas se asignan de manera aleatoria.

Finalmente se recomienda considerar las notas de preguntas que no sean de selección múltiple o la inclusión de las mismas, y analizar complementariamente que implicaciones tienen sobre el rendimiento de los estudiantes y sus comportamientos con sospecha de fraude, estos sobre diferentes casos de estudio, ya que este tipo de notas no se tuvieron en cuenta.

# Bibliography

- [1] D. Jaramillo-Morillo, J. Ruipérez-Valiente, M. F. Sarasty, and G. Ramírez-Gonzalez, “Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics,” vol. 17, no. 1, p. 45. [Online]. Available: <https://doi.org/10.1186/s41239-020-00221-2>
- [2] Open source LMSs continue to advance with open edX. [Online]. Available: <http://www.iacet.org/news/iacet-blog/blog-articles/open-source-lmss-continue-to-advance-with-open-edx/>
- [3] D. Jaramillo-Morillo, M. S. Sarasty, and G. Ram, “Follow-up of learning activities in open edX: A case study at the university of cauca,” p. 6.
- [4] C. Almenara, A. Isabel, V. Martínez, and U. de Sevilla, “LAS TIPOLOGÍAS DE MOOC: SU DISEÑO E IMPLICACIONES EDUCATIVAS,” p. 14.
- [5] L. García Aretio, “MOOC: ¿tsunami, revolución o moda pasajera?” *RIED. Revista Iberoamericana de Educación a Distancia*, vol. 18, no. 1, Dec. 2014. [Online]. Available: <http://revistas.uned.es/index.php/ried/article/view/13812>
- [6] W. Jobe, “No university credit, no problem? Exploring recognition of non-formal learning,” in *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*. Madrid, Spain: IEEE, Oct. 2014, pp. 1–7. [Online]. Available: <http://ieeexplore.ieee.org/document/7044389/>
- [7] —, “No university credit, no problem? Exploring recognition of non-formal learning,” in *2014 IEEE Frontiers in Education Conference (FIE)*

- Proceedings*. Madrid, Spain: IEEE, Oct. 2014, pp. 1–7. [Online]. Available: <http://ieeexplore.ieee.org/document/7044389/>
- [8] “Learn the Latest Tech Skills; Advance Your Career | Udacity.” [Online]. Available: <https://www.udacity.com>
- [9] “Coursera | Online Courses & Credentials by Top Educators. Join for Free.” [Online]. Available: <https://www.coursera.org/>
- [10] “edX | Cursos online gratis de Harvard, MIT y más.” [Online]. Available: <https://www.edx.org/es>
- [11] X. Sun, W. Zhou, Q. Xiang, B. Cui, and Y. Jin, “Research on big data analytics technology of MOOC,” in *2016 11th International Conference on Computer Science & Education (ICCSE)*. IEEE, pp. 64–68. [Online]. Available: <http://ieeexplore.ieee.org/document/7581556/>
- [12] R. M. Paton, A. E. Fluck, and J. D. Scanlan, “Engagement and retention in VET MOOCs and online courses: A systematic review of literature from 2013 to 2017,” *Computers & Education*, vol. 125, pp. 191–201, Oct. 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360131518301568>
- [13] R. Pollack Ichou, “Can MOOCs reduce global inequality in education?” *Australasian Marketing Journal (AMJ)*, vol. 26, no. 2, pp. 116–120, May 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1441358218301083>
- [14] “MOOCs 2018: A Year in Review Class Central.” [Online]. Available: <https://www.classcentral.com/moocs-year-in-review-2018>
- [15] S. Suresh Kumar and P. M. Mallikarjuna Shastry, “Analysis of student engagement and course completion in massive open online courses,” in *Integrated Intelligent Computing, Communication and Security*, A. Krishna, K. Srikantaiah, and C. Naveena, Eds. Springer Singapore, vol. 771, pp. 447–458. [Online]. Available: [http://link.springer.com/10.1007/978-981-10-8797-4\\_46](http://link.springer.com/10.1007/978-981-10-8797-4_46)

- [16] I. Aguaded and R. Medina-Salguero, “Certificación de los MOOC y su reconocimiento en créditos universitarios,” p. 12.
- [17] “¿Qué son los programas MicroMasters™?” [Online]. Available: <http://support.edx.org/hc/es-419/articles/227511108--Qu%C3%A9-son-los-programas-MicroMasters->
- [18] D. Jaramillo-Morillo, M. Solarte Sarasty, G. Ramírez González, and M. Pérez-Sanagustín, “Follow-up of learning activities in open edX: A case study at the university of cauca,” in *Digital Education: Out to the World and Back to the Campus*, C. Delgado Kloos, P. Jermann, M. Pérez-Sanagustín, D. T. Seaton, and S. White, Eds. Springer International Publishing, vol. 10254, pp. 217–222. [Online]. Available: [http://link.springer.com/10.1007/978-3-319-59044-8\\_26](http://link.springer.com/10.1007/978-3-319-59044-8_26)
- [19] P. M. Moreno-Marcos, P. J. Muñoz-Merino, C. Alario-Hoyos, I. Estévez-Ayres, and C. Delgado Kloos, “Analysing the predictive power for anticipating assignment grades in a massive open online course,” *Behaviour & Information Technology*, vol. 37, no. 10-11, pp. 1021–1036, Nov. 2018. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/0144929X.2018.1458904>
- [20] G. Alexandron, J. A. Ruipérez-Valiente, Z. Chen, P. J. Muñoz-Merino, and D. E. Pritchard, “Copying@scale: Using harvesting accounts for collecting correct answers in a MOOC,” vol. 108, pp. 96–114. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360131517300234>
- [21] J. Peral, A. Maté, and M. Marco, “Application of Data Mining techniques to identify relevant Key Performance Indicators,” *Computer Standards & Interfaces*, vol. 50, pp. 55–64, Feb. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0920548916300927>
- [22] J.-S. Wong and X. L. Zhang, “Messagelens : A visual analytics system to support multifaceted exploration of mooc forum discussions,” *Visual Informatics*, vol. 2, no. 1, pp. 37–49, Mar. 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2468502X18300111>

- [23] H. Qu and Q. Chen, "Visual Analytics for MOOC Data," *IEEE Computer Graphics and Applications*, vol. 35, no. 6, pp. 69–75, Nov. 2015. [Online]. Available: <http://ieeexplore.ieee.org/document/7331178/>
- [24] C. Shi, S. Fu, Q. Chen, and H. Qu, "VisMOOC: Visualizing Video Clickstream Data from Massive Open Online Courses," p. 2.
- [25] S. Fu, J. Zhao, W. Cui, and H. Qu, "Visual Analysis of MOOC Forums with iForum," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 201–210, Jan. 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/7534878/>
- [26] M. Klusener and A. Fortenbacher, "Predicting students' success based on forum activities in MOOCs," in *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. Warsaw, Poland: IEEE, Sep. 2015, pp. 925–928. [Online]. Available: <http://ieeexplore.ieee.org/document/7341439/>
- [27] S. Fu, J. Zhao, H. F. Cheng, H. Zhu, and J. Marlow, "T-Cal: Understanding Team Conversational Data with Calendar-based Visualization," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. Montreal QC, Canada: ACM Press, 2018, pp. 1–13. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=3173574.3174074>
- [28] Q. Chen, Y. Chen, D. Liu, C. Shi, Y. Wu, and H. Qu, "PeakVizor: Visual Analytics of Peaks in Video Clickstreams from Massive Open Online Courses," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 10, pp. 2315–2330, Oct. 2016. [Online]. Available: <http://ieeexplore.ieee.org/document/7346501/>
- [29] "PMBOK Guide and Standards | Project Management Institute." [Online]. Available: <https://www.pmi.org/pmbok-guide-standards>
- [30] M. H. Baturay, "An overview of the world of MOOCs," vol. 174, pp. 427–433. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1877042815007363>

- [31] —, “An overview of the world of moocs,” *Procedia - Social and Behavioral Sciences*, vol. 174, pp. 427–433, 02 2015.
- [32] P. R. Martín and C. M. T. Artigas, “ANÁLISIS DE LA OFERTA COMPLETA DE CURSOS DE LAS PLATAFORMAS COURSERA, EdX, MIRÍADA x y UDACITY,” p. 95.
- [33] V. Mailhes and J. Raspa, “Mooc: De la revolución educativa a la supervivencia,” p. 17.
- [34] E. Tovar, R. Cabedo, A. Alzaghoul, S. Martin, M. Llamas, M. Caeiro, O. M. Bonastre, R. Strachan, and M. Castro, “The MOOC offer in open education: A revision of the first IEEE education society’ MOOC,” in *2018 Learning With MOOCS (LWMOOCS)*. IEEE, pp. 151–154. [Online]. Available: <https://ieeexplore.ieee.org/document/8534675/>
- [35] F. J. Benito Justel, R. F. Beijinho do Rosário, L. D. Ben Abdellah, and J. F. Ávila de Tomás, “¿qué son los cursos abiertos en línea y masivos? nuevas formas de aprender sin horarios ni fronteras,” vol. 21, no. 7, pp. 424–429. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1134207214708059>
- [36] F. J. Palacios Hidalgo, C. A. Huertas Abril, and M. E. Gómez Parra, “MOOCs: Origins, concept and didactic applications: A systematic review of the literature (2012–2019),” vol. 25, no. 4, pp. 853–879. [Online]. Available: <http://link.springer.com/10.1007/s10758-019-09433-6>
- [37] V. F. de Moura, C. A. de Souza, and A. B. N. Viana, “The use of massive open online courses (MOOCs) in blended learning courses and the functional value perceived by students,” p. 14.
- [38] K. Julia, V. R. Peter, and K. Marco, “Educational scalability in MOOCs: Analysing instructional designs to find best practices,” vol. 161, p. 104054. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360131520302529>

- [39] R. Mahajan, P. Gupta, and T. Singh, “Massive open online courses: Concept and implications,” vol. 56, no. 6, pp. 489–495. [Online]. Available: <http://link.springer.com/10.1007/s13312-019-1575-6>
- [40] A. López de la Serna, C. Castaño Garrido, and D. Herrero Fernández, “Integración de los cursos SPOC en las asignaturas de grado. una experiencia práctica,” no. 52, pp. 139–149. [Online]. Available: <https://recyt.fecyt.es/index.php/pixel/article/view/62528>
- [41] A. M. Kaplan and M. Haenlein, “Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the cookie monster,” vol. 59, no. 4, pp. 441–450. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S000768131630009X>
- [42] K. Piersig, M. Egloffstein, M. Pumpat, A. Eckhardt, and D. Wagner, “Designing SPOCs for student engagement – experiences from management education,” p. 7.
- [43] R. A. Sarasa and J. Bravo-Agapito, “Implantación de un SPOC en la educación a distancia para la mejora del proceso de enseñanza-aprendizaje,” vol. 0, no. 6, number: 6. [Online]. Available: <https://tecnologia-ciencia-educacion.com/index.php/TCE/article/view/119>
- [44] E. C. Magaña, E. S. Rivas, J. M. F. Lacorte, and J. M. T. Torres, “SPOC y formación del profesorado: aproximación bibliométrica y pedagógica en scopus y web of science,” vol. 23, no. 2, number: 2. [Online]. Available: <https://revistas.um.es/reifop/article/view/413541>
- [45] D.-S. Chen and G.-P. Yan, “MPOC experiential learning model of engineering education based on distributed flip in local engineering colleges and universities in “post MOOC” era,” vol. 20, no. 6, pp. 1375–1379. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/09720529.2017.1392449>
- [46] W. Guo, “From spoc to mpoc – the effective practice of peking university online teacher training,” in *2014 International Conference of Educational Innovation through Technology*, Oct 2014, pp. 258–264.

- [47] H. M. Dai, T. Teo, and N. A. Rappa, "Understanding continuance intention among MOOC participants: The role of habit and MOOC performance," vol. 112, p. 106455. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0747563220302077>
- [48] A. Mali, "Massive open online courses," vol. 3, pp. 1–9.
- [49] C. Baker, F. Nafukho, K. McCaleb, M. Becker, and M. Johnson, "The tangible and intangible benefits of offering massive open online courses: Faculty perspectives," vol. 4, no. 2. [Online]. Available: <http://www.ipsonet.org/publications/open-access/internet-learning/volume-4-number-2-fall-2015>
- [50] Y. Aljaraideh, "Massive open online learning (MOOC) benefits and challenges: A case study in Jordanian context," vol. 12, pp. 65–78.
- [51] J. Fang, L. Tang, J. Yang, and M. Peng, "Social interaction in MOOCs: The mediating effects of immersive experience and psychological needs satisfaction," vol. 39, pp. 75–91. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0736585318308906>
- [52] D. Carrizo and C. Moller, "Estructuras metodológicas de revisiones sistemáticas de literatura en ingeniería de software: un estudio de mapeo sistemático," vol. 26, pp. 45–54. [Online]. Available: [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-33052018000500045&lng=en&nrm=iso&tlng=en](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-33052018000500045&lng=en&nrm=iso&tlng=en)
- [53] K. Sarajeva and SPIDER, *ICT for anti-corruption, democracy and education in East Africa*. Kista: SPIDER, 2013, oCLC: 941331481.
- [54] P. McGee, "Supporting Academic Honesty in Online Courses," *The Journal of Educators Online*, vol. 10, no. 1, Jan. 2013. [Online]. Available: [https://www.thejeo.com/archive/2013\\_10\\_1/mcgee](https://www.thejeo.com/archive/2013_10_1/mcgee)
- [55] B. J. Zirger, E. Rutz, D. Boyd, J. Tappel, and V. Subbian, "Creating pathways to higher education: A cross-disciplinary MOOC with graduate credit," in *2014 IEEE Integrated STEM Education Conference*. Princeton, NJ, USA: IEEE, Mar. 2014, pp. 1–5. [Online]. Available: <http://ieeexplore.ieee.org/document/6891056/>



- [56] C. Gherib, J.-M. Dujardin, and D. Verpoorten, "MOOCS IN BUSINESS ADMINISTRATION – AN OVERVIEW OF ASSESSMENT PRACTICE," p. 15.
- [57] T. Staubitz, R. Teusner, J. Renz, and C. Meinel, "An Experiment in Automated Proctoring," p. 12, 2016.
- [58] J. Miguel, S. Caballe, and J. Prieto, "Providing Information Security to MOOC: Towards Effective Student Authentication," in *2013 5th International Conference on Intelligent Networking and Collaborative Systems*. Xi'an city, Shaanxi province, China: IEEE, Sep. 2013, pp. 289–292. [Online]. Available: <http://ieeexplore.ieee.org/document/6630424/>
- [59] R. Woolsey, "Student verification system for online assessments: Bolstering quality and integrity of distance learning," *Journal of Industrial Technology*, vol. Volume 27, 09 2011.
- [60] Y. Atoum, L. Chen, A. X. Liu, S. D. H. Hsu, and X. Liu, "Automated Online Exam Proctoring," *IEEE Transactions on Multimedia*, vol. 19, no. 7, pp. 1609–1624, Jul. 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/7828141/>
- [61] S. Prathish, Athi Narayanan S, and K. Bijlani, "An intelligent system for online exam monitoring," in *2016 International Conference on Information Science (ICIS)*. Kochi, India: IEEE, Aug. 2016, pp. 138–143. [Online]. Available: <http://ieeexplore.ieee.org/document/7845315/>
- [62] M. Cote, F. Jean, A. B. Albu, and D. Capson, "Video summarization for remote invigilation of online exams," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Lake Placid, NY: IEEE, Mar. 2016, pp. 1–9. [Online]. Available: <http://ieeexplore.ieee.org/document/7477704/>
- [63] E. G. dice, "Coursera introduce 'Signature Track', una forma de obtener un certificado que verifica la identidad del alumno." Jan. 2013. [Online]. Available: [https://pilartarrino.wordpress.com/2013/01/11/coursera-introduce-signatura-track-una-forma-de-obtener-un-certificado-que-verifica-la-identidad](https://pilartarrino.wordpress.com/2013/01/11/coursera-introduce-signatura-track-una-forma-de-obtener-un-certificado-que-verifica-la-identidad/)

- [64] C. G. Northcutt, A. D. Ho, and I. L. Chuang, “Detecting and preventing “multiple-account” cheating in massive open online courses,” *Computers & Education*, vol. 100, pp. 71–80, Sep. 2016. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360131516300896>
- [65] J. A. Ruiperez-Valiente, P. J. Munoz-Merino, G. Alexandron, and D. E. Pritchard, “Using Machine Learning to Detect ‘Multiple-Account’ Cheating and Analyze the Influence of Student and Problem Features,” *IEEE Transactions on Learning Technologies*, pp. 1–1, 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/8219749/>
- [66] J. A. Ruipérez-Valiente, S. Joksimović, V. Kovanović, D. Gašević, P. J. Muñoz-Merino, and C. Delgado Kloos, “A Data-driven Method for the Detection of Close Submitters in Online Learning Environments,” in *Proceedings of the 26th International Conference on World Wide Web Companion - WWW '17 Companion*. Perth, Australia: ACM Press, 2017, pp. 361–368. [Online]. Available: <http://dl.acm.org/citation.cfm?doi=3041021.3054161>
- [67] A. Gandomi and M. Haider, “Beyond the hype: Big data concepts, methods, and analytics,” vol. 35, no. 2, pp. 137–144. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0268401214001066>
- [68] G. Korosi and F. Havasi, “Moodle-based data mining potentials of MOOC systems at the University of Szeged,” in *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. Opatija, Croatia: IEEE, May 2017, pp. 755–760. [Online]. Available: <http://ieeexplore.ieee.org/document/7973523/>
- [69] N. Hakami, S. White, and S. Chakaveh, “Motivational Factors that Influence the use of MOOCs: Learners’ Perspectives - A Systematic Literature Review:,” in *Proceedings of the 9th International Conference on Computer Supported Education*. Porto, Portugal: SCITEPRESS - Science and Technology Publications, 2017, pp. 323–331. [Online]. Available: <http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006259503230331>

- [70] R. Pérez-Álvarez, J. Maldonado-Mahauad, and M. Pérez-Sanagustín, “How to map learning activities through URLs? The case of Coursera platform,” p. 10, 2018.
- [71] D. Keim, “Information visualization and visual data mining,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 1, pp. 1–8, Mar. 2002. [Online]. Available: <http://ieeexplore.ieee.org/document/981847/>
- [72] Q. Chen, X. Yue, X. Plantaz, Y. Chen, C. Shi, T.-C. Pong, and H. Qu, “Viseq: Visual analytics of learning sequence in massive open online courses,” *IEEE transactions on visualization and computer graphics*, vol. PP, 10 2018.
- [73] X. Li, C. Men, F. Zhang, and Z. Du, “A Smart Visual Analysis Solution for MOOC Data,” in *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*. Orlando, FL: IEEE, Nov. 2017, pp. 101–106. [Online]. Available: <http://ieeexplore.ieee.org/document/8328374/>
- [74] D. Jaramillo-Morillo, M. Solarte, and G. R. González, “Estrategia de seguimiento a las actividades de aprendizaje de los estudiantes en cursos en línea masivos y privados (MPOC) con reconocimiento académico en la universidad del cauca,” p. 20.
- [75] R. A. P. Morocho and J. A. V. Mendoza, “Descubrimiento de patrones de interacción en cursos mooc en entornos online: un enfoque utilizando minería de procesos. caso de estudio: “curso de la metodología dicrevoa en open edx”,” 2016.
- [76] D. A. J. Morillo, “SEGUIMIENTO AUTOMÁTICO DE ACTIVIDADES DE APRENDIZAJE EN CURSOS EN LÍNEA ABIERTOS y MASIVOS,” p. 121.
- [77] The platform – open edX. [Online]. Available: <https://open.edx.org/the-platform/>

- [78] M. J. L. Lapuente and C. L. Lapuente. HTML. Publisher: Lamarca Lapuente. [Online]. Available: <http://www.hipertexto.info/documentos/html.htm>
- [79] Q. Devs. ¿qué es angular y para qué sirve? [Online]. Available: <https://www.qualitydevs.com/2019/09/16/que-es-angular-y-para-que-sirve/>
- [80] P. Kruchten, “Planos arquitectónicos: El modelo de “4+1” vistas de la arquitectura del software,” p. 16.
- [81] Unify the DevOps lifecycle with GitLab. [Online]. Available: <https://about.gitlab.com/stages-devops-lifecycle/>
- [82] R. Swaray, “An evaluation of a group project designed to reduce free-riding and promote active learning,” *Assessment & Evaluation in Higher Education*, vol. 37, no. 3, pp. 285–292, 2012, publisher: Routledge  
\_eprint: <https://doi.org/10.1080/02602938.2010.531246>. [Online]. Available: <https://doi.org/10.1080/02602938.2010.531246>

# Anexos

# Anexo A

## Manual de Usuario SeleneVis

### A.1 Vista login

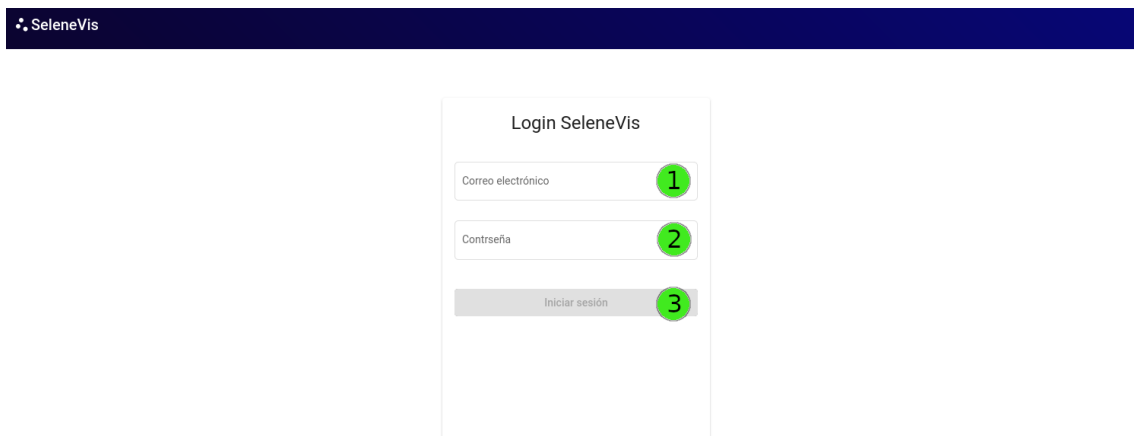


Figura A.1 Vista de login.

1. **Correo Electrónico:** en este campo debe ir el correo del usuario con el que se registró en la aplicación.
2. **Contraseña:** en este campo se debe escribir la contraseña asignada al usuario con el correo correspondiente. Estas credenciales las provee el administrador de la plataforma.

3. **Botón iniciar sesión:** Después de ingresar las credenciales de usuario, este botón se activa para entrar a la aplicación.

## A.2 Vista estadísticas

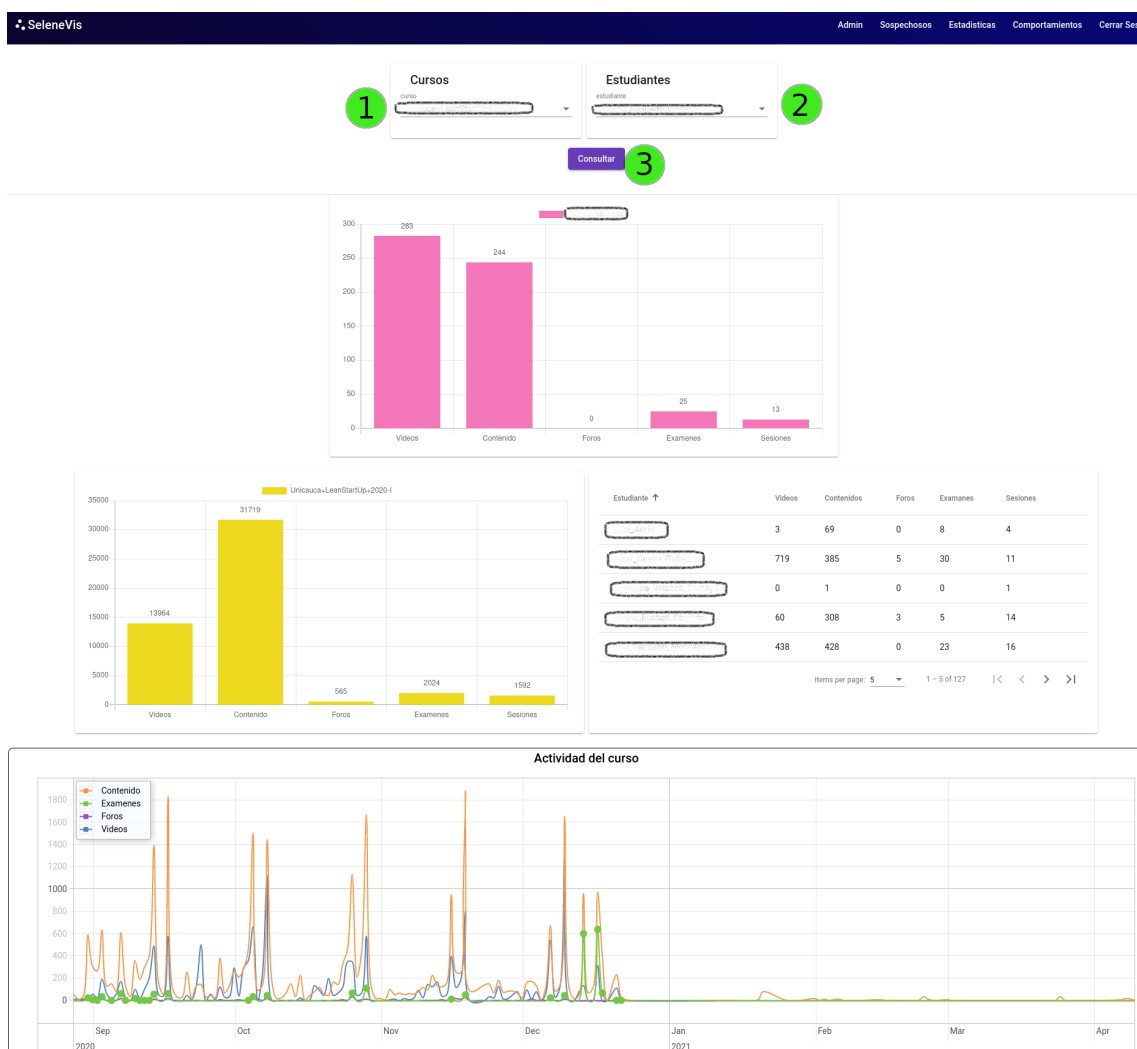


Figura A.2 Vista estadísticas.

1. **Lista de cursos:** en esta lista de cursos, se debe seleccionar el curso que se desee analizar. Es importante tener en cuenta que los cursos desplegados en la

lista son los asignados por el administrador al docente. Una vez seleccionado el curso, se presentan las estadísticas generales del curso mediante gráficas.

2. **Lista de estudiantes:** corresponde a la lista de estudiantes que están inscritos en el curso seleccionado en el punto 1.
3. **Botón “consultar”:** si se requiere, puede ser seleccionado de la lista el nombre de un estudiante y proceder a pulsar el botón de consultar para obtener las estadísticas generales del estudiante en ese curso específico.

## A.3 Vista sospechosos

Nombre	Fecha de Presentación	# Preguntas del Examen	Coincidencias	Diferencia de Tiempo	Procentaje similitud
Sospechoso 5	10/29/2020, 4:57:49 PM	9	9	0H:0M:3s	100%
Sospechoso 6	10/29/2020, 4:57:46 PM				

Figura A.3 Vista sospechosos.

- **Lista cursos:** en esta lista de cursos, se debe seleccionar el curso que se desee analizar. Es importante tener en cuenta que los cursos desplegados en la lista son los asignados por el administrador al docente.
- **Lista exámenes:** en esta lista aparecen los exámenes que se desarrollaron a lo largo del curso seleccionado, se debe seleccionar el examen el cual se quiere analizar.
- **Lista sospechosos:** se debe seleccionar los dos estudiantes a comparar, esta lista de estudiantes hace referencia a los estudiantes sospechosos en el examen del curso que se seleccionaron anteriormente.



- **Botón “consultar”**: se debe pulsar el botón para poder ver en pantalla la comparación de los estudiantes seleccionados.

## A.4 Vista comportamientos



Figura A.4 Vista comportamientos.

1. **Lista cursos**: en esta lista de cursos, se debe seleccionar el curso que se desee analizar. Es importante tener en cuenta que los cursos desplegados en la lista son los asignados por el administrador al docente.
2. **Lista estudiantes**: se debe seleccionar al menos un estudiante con el fin de ver

su comportamiento, esta lista de estudiantes hace referencia a los alumnos del curso seleccionado, una vez seleccionado un estudiante se presenta en pantalla, las gráficas de comportamiento del mismo.

3. **Opciones de consulta:** existen dos opciones; “consulta por día” y “consulta por sesión”, por defecto aparece seleccionada la primera opción nombrada, la cual permite observar el comportamiento del estudiante a través de los días en que tuvo interacción con la plataforma dentro del curso seleccionado, la segunda opción de “consulta por sesión”, permite observar el comportamiento del estudiante a través de las diferentes sesiones que se detectaron dentro del curso.
4. **Lista de opción:** en esta lista aparecen los diferentes días o sesiones del estudiante, dependiendo de la opción seleccionada en el punto 3, se debe seleccionar una opción y se desplegará en pantalla las gráficas de comportamiento de la opción en particular.

## A.5 Vista administración

The screenshot displays the administration interface. At the top right, there is a navigation bar with the following items: Admin, Sospechosos, Estadísticas, and Comportamientos. On the left side, there is a form titled "Ingresar Docente" (1) with the following fields: Nombre, Curso, Correo, Seleccione un perfil, Contraseña, Repetir Contraseña, and a "Guardar" button (2). The main area shows a table with the following columns: Nombre, Correo, Cursos, Perfil, and Operacion. The table contains several rows of user data, with the first row highlighted and a green circle (3) next to the "Operacion" column.

Nombre	Correo	Cursos	Perfil	Operacion
[Redacted]	[Redacted]	10 cursos	Administrador	[Pencil icon] [Trash icon] (3)
[Redacted]	[Redacted]	1 cursos	Docente	[Pencil icon] [Trash icon]
[Redacted]	[Redacted]	2 cursos	Administrador	[Pencil icon] [Trash icon]
[Redacted]	[Redacted]	1 cursos	Docente	[Pencil icon] [Trash icon]
[Redacted]	[Redacted]	1 cursos	Docente	[Pencil icon] [Trash icon]
[Redacted]	[Redacted]	2 cursos	Docente	[Pencil icon] [Trash icon]
[Redacted]	[Redacted]	1 cursos	Administrador	[Pencil icon] [Trash icon]
[Redacted]	[Redacted]	3 cursos	Administrador	[Pencil icon] [Trash icon]

Figura A.5 Vista administración.

1. **Formulario docente:** se debe ingresar todos los datos del formulario para almacenar un nuevo docente en el sistema, como lo es el nombre, los cursos a los cuales dicta clase (este campo es una lista de cursos de los cuales se puede seleccionar más de uno), el correo electrónico, el perfil de docente (administrador o solo docente) y finalmente la contraseña para que el docente puede entrar al sistema.
2. **Botón guardar:** este botón se habilita una vez el formulario esté completo, es necesario accionarlo una vez se esté seguro de que los datos del docente a almacenar sean correctos.
3. **Opciones de editar y eliminar:** la opción de editar se usa para corregir algún dato de un registro ya guardado, se puede editar todos los campos del formulario del punto 1. La opción de eliminar, borra completamente el registro de la base de datos, antes de ser eliminado deberá mostrar un mensaje preguntando si en realidad se desea eliminar dicho registro.

# Anexo B

## Manual de instalación Selenevis

### B.1 Dependencias

Para un buen funcionamiento de la aplicación es necesario tener instalados los siguientes requerimientos:

1. **Node js**: <https://nodejs.org/es/> en cualquier versión.
2. **Angular cli**: se instala con el comando “npm install -g @angular/cli”.
3. **MongoDB community**: <https://www.mongodb.com/try/download/community>.

### B.2 Clonar repositorio

En primera medida se debe clonar o descargar el repositorio de la aplicación, disponible en <https://gitlab.com/aldairl/SeleneVis.git>.

## B.3 Instalar dependencias

1. **Dependencias backend:** una vez descargado o clonado el repositorio, se procede a instalar las dependencias de la misma, para ello, se abre una consola dentro de la carpeta de la aplicación donde se encuentre el archivo **“package.json”** y se procede a ejecutar el comando **“npm install”**, este comando instalará todas las dependencias necesarias para el funcionamiento de la aplicación a nivel de backend.
2. **Dependencias frontend:** Después de haber instalado las dependencias del backend, se procede a instalar las dependencias del frontend, es decir de la parte visual de la aplicación, para ello se ingresa a la carpeta **“frontend”** a través de la consola se ejecuta el comando **“npm install”**, dicho comando instalará todas las dependencias de la aplicación.

## B.4 Generar archivos producción

Una vez instaladas todas las dependencias se procede a generar los archivos de producción de la aplicación, es decir los archivos los cuales se van a subir a un servidor web o para probar en local, para ello se ejecuta el comando **“npm run builddeploy”**.

## B.5 Configurar base de datos

Es necesario crear la base de datos llamada **“seguimiento”** y guardar dentro de la colección docentes el administrador de la plataforma para poder iniciar sesión, para ello se abre el shell de mongo con el comando **“mongo”**.

1. **Crear base de datos:** Una vez dentro del shell de mongo solo basta con usar el comando **“use seguimiento”**.

2. **Crear administrador:** Se crea con la sentencia mostrada en la figura [B.1](#).

```
db.docentes.insert({"course":["Unicauca+C_Enfermeria+2019-II"],"name":"admin-na  
me","correo":"admin-correo@gmail.com","password":"la-password-del-admin-encrypt  
ada","credencial":"Administrador"})
```

Figura B.1 Elaboración propia. Crear usuario administrador.

## B.6 Lanzar aplicación

Con todos los pasos anteriores, es posible lanzar la aplicación, para ello se abre una consola en la carpeta del proyecto y se procede a ejecutar la instrucción “**node server/index.js**”, en consola deberá aparecer el puerto donde se está ejecutando la aplicación, por defecto corre en el puerto 8080. Se abre el navegador en la dirección “**http://localhost:8080**” y aparecerá la vista de login.

**Anexo C**

**Artículo**

# Visualización científica de los comportamientos con sospecha de fraude en cursos en línea, abiertos y masivos MOOC

Aldair Luna Buitrón, Herman Gabriel Eraso R.

**Abstract**—MOOCs (Massive Open Online Courses) have caused great impact in the academic world, creating a new way to access education in a scalable, flexible and asynchronous way. They are currently being implemented and academically accredited through SPOC (Small Private Online Course) and MPOC (Massive Private Online Course) variants within universities and higher education institutions. However, MOOCs were not initially designed to be certified and recognized, so they do not take into account possible cases of low performance, learning difficulties that generate high dropout rates and behaviors with suspected fraud, which have been identified in courses that present recognition or validation of educational credits. In addition, the different MOOC platforms do not offer the necessary tools to support teachers, monitor students and detect suspected fraudulent behavior. The objective of this work is to build a visualization mechanism through a web development "SELENEVIS", which represents a support tool for the teacher and allows him to visualize the students identified as suspects and explore the record of their activities. As a result, there are 4 views that represent different levels of abstraction. An Administrator view, a suspect identification view supported by the work[1], a per-student or per-course statistics view, and a comparison view that allows comparison of activities between students using a proposed graph scheme of activity sequences in conjunction with complementary statistical plots.

**Keywords**—Mooc, Academic dishonesty, Visualization, Suspect identification, Follow-up, Graphs schema.

## I. INTRODUCCIÓN

Los Cursos en Línea Masivos y Abiertos (MOOC, Massive Open Online Courses) abarcan gran cantidad de estudiantes (cientos o miles de participantes) y que toman diferentes áreas de la educación, desde la poesía hasta la tecnología, son cursos en línea tradicionales, pero se diferencian por su carácter abierto y masivo [2].

La característica de ser abierto implica que los recursos educativos (por lo general basado en multimedia) están al alcance de todos y flexibilizan el proceso de formación en acceso y tiempo. La masividad por su parte consiste en que no existe un límite de inscritos en un curso, por lo que este puede contar con la participación de miles de estudiantes [3].

En la actualidad existe una gran variedad de plataformas MOOC, en cada una de ellas se generan gran cantidad de datos, entre estos se encuentran los rastros de las actividades que los estudiantes realizan dentro del curso, estos datos han sido usados con el fin de estudiar los cursos MOOC y el aprendizaje en ellos, brinda a los investigadores de la educación, información valiosa para entender y mejorar los cursos MOOC [4].

Los MOOC se dividen en dos grandes grupos: los SPOC y los MPOC. Los SPOC (Small Private Online Course) no

es de carácter masivo, sino que su número de participantes alcanza unas pocas decenas o hasta unos cientos de estudiantes, es privado, por tanto el acceso no es completamente abierto a todo el mundo, sino a un grupo selecto de personas [5].

Los MPOC (Massive Private Online Course), se distinguen por estar restringidos a través de una característica de privacidad, es decir los MPOC no están completamente abiertos a todo público ni son necesariamente gratuitos, por lo que el número de estudiantes inscritos no es tan grande como en los MOOC, pero si suelen tomar números mayores de participantes en contraste con la cantidad de aprendices inscritos en los cursos en línea convencionales o los SPOC [6].

Los MOOC han promovido enfoques de analítica y visualización, donde la visualización científica desarrolla un papel importante a la hora de adaptarse a nuevas formas de visualización para poder mostrar las características, patrones e información de la forma más adecuada. Este trabajo presenta un herramienta de visualización científica enmarcada en el contexto de los cursos en línea abiertos y masivos MOOC y cuyo aporte permite establecer una presentación visual de los comportamientos con sospecha de fraude.

El documento presenta los trabajos relacionados en la sección II, la descripción del caso de estudio en la sección III, se describe y analiza el funcionamiento de la Herramienta de visualización "SELENEVIS" en la sección IV, y finalmente se plantean conclusiones y trabajos futuros en la sección V.

## II. TRABAJOS RELACIONADOS

### A. Visualización

Existen diferentes trabajos centrados en la visualización que permiten observar de una manera mas clara los datos de plataformas MOOC, entre ellos, [7] una herramienta llamada "VisMooc", La cual permite el análisis del flujo de clics sobre los vídeos del curso, análisis de foros y discutir nuevos hallazgos sobre los comportamientos de aprendizaje. [8] se enfoca en el foro del curso con una herramienta llamada "Mensaje Lens", que permite un análisis visual de los datos del foro como lo es; saber el nivel de similitud de un tema en específico, saber cuáles son las palabras más usadas y brinda una representación gráfica de la red de interacción entre estudiantes dentro del foro en diferentes vistas. [4] "MoocViz" un framework multiplataforma para la investigación de la educación en MOOC, entre sus herramientas permite conocer el numero de estudiantes que obtuvieron un certificado después de finalizar un curso, permite saber en número de estudiantes matriculados en un país determinado



y visualizar comportamientos como por ejemplo, que hay un aumento de ingresos a la plataforma justo antes de finalizar una actividad del curso, [9] tienen una solución de análisis visual inteligente para datos MOOC, en esta solución permiten observar el progreso del estudiante, crearon 4 vistas que las llamaron, 'vista demográfica', 'vista de actividades de calendario', 'vista de distribución de progreso' y 'vista de rastro personal' dentro de un curso. Sin embargo no dan un enfoque hacia la deshonestidad académica.

### B. Reconocimiento Académico

Existen varios intentos por validar estrategias de reconocimiento académico. [10] intenta validar el fortalecimiento e interés del estudiantes con el curso a partir de la inclusión de incentivos e insignias, sin tener en cuenta casos de deshonestidad académica ni estrategias de visualización. [2] plantea estrategias de reconocimiento académico a partir de credenciales digitales y la influencia de las mismas en el comportamiento de los estudiantes, sin embargo no valida la deshonestidad académica y toda la retroalimentación es a partir de entrevistas y encuestas. en [11] por otra parte se implementa estrategias de reconocimiento académico a partir de la validación de créditos educativos dentro de la Universidad para aquellos estudiantes que apruebe el curso MPOC dispuesto por la misma institución.

### C. Deshonestidad académica

Esfuerzos para resolver la preocupación de deshonestidad académica han llevado a utilizar estrategias que la plantean la necesidad de supervisión. [12], [13], [14] plantean estrategias de supervisión en las que acarrea la necesidad de exámenes presenciales. [15], [16], [17] plantean una estrategia de supervisión automática en la que se utilizan estrategias de reconocimiento facial, revisión de la integridad a partir de cámaras web, el uso de micrófonos o reconocimiento biométrico. En los cuales se requiere de la presencia y revisión de un examinador final. En [18], [19], [20] trabajan la supervisión automática en línea. Por otra parte, un enfoque diferente se presenta a partir del análisis del aprendizaje. [40,41,42,43] plantean el análisis de conductas inapropiadas a través de comportamiento de los estudiantes con la plataforma. Se tiene entonces información estadística de casos de CAMEO y de vínculos de participación invisibles. Esto sin tener en cuenta una herramienta de visualización que permita observar el comportamiento con sospecha de fraude de los estudiantes a partir de su interacción con la plataforma.

## III. CASO DE ESTUDIO

En el caso de estudio se ha realizado una descripción del curso, las estrategias propuestas para el mismo y la disposición de las actividades de evaluación de los estudiantes con sus respectivos porcentajes de calificación. Por otra parte, a partir de esta información se obtendrán los resultados tomados a partir del uso de la aplicación Web SelenVis.

### A. Descripción del curso

Para este caso de estudio se ha tomado de referencia el curso "Introducción al emprendimiento con Lean Startup", este curso habilitado el 24 de Agosto de 2020 y con fecha de finalización el día 21 de Diciembre del 2020.

El curso tiene como propósito fundamental, introducir a los estudiantes en el conocimiento y uso de una de las metodologías para el desarrollo de emprendimientos que mayor éxito ha tenido en los últimos tiempos: el Lean Startup.

Al ser el curso diseñado desde una perspectiva no presencial, se basa en el desarrollo de actividades de aprendizaje semanales soportadas en las Tecnologías de Internet, donde el estudiante tendrá acceso a diversos recursos didácticos construidos y seleccionados por el profesor, independiente de otros que bien pueda conseguir y consultar, además de la participación en foros de discusión, la elaboración de mapas conceptuales, uso de simuladores en línea, la presentación de pruebas en línea y el desarrollo de talleres no presenciales.

Durante la primera semana de clase se permite la interacción con los servicios de la plataforma tecnológica que soporta el curso, para que los estudiantes se familiaricen con las funciones y actividades que van a desarrollar durante el curso. Posteriormente, semana a semana, los estudiantes tendrán acceso a un conjunto de recursos digitales que deberán consultar previo al desarrollo de actividades de aprendizaje, algunas de ellas pueden ser calificables y tenerse en cuenta en la obtención de la evaluación del estudiante a lo largo del curso.

El curso fue dispuesto con dos unidades temáticas que conforman un total de 5 módulos que se habilitan semana a semana, de tal forma que en total se desarrollan 5 exámenes correspondientes a cada uno de ellos. Por otra parte, se tiene un examen final que evalúa el total de unidades del curso.

Los exámenes se habilitan 14 días luego de haberse dispuesto el material de cada una de las unidades. Además los estudiantes tienen la posibilidad de presentar un supletorio dos días después del examen oficial.

El examen final dispuesto en el curso, no es el mismo para todos los estudiantes, dado que se ha asignado un banco de preguntas, de las cuales, se asignan de forma aleatoria para formar un examen diferente a cada uno de ellos.

algunos de los exámenes se componen de dos partes, las preguntas de selección múltiple y actividades que las califica el docente encargado y asigna una respectiva nota, Para el análisis de este trabajo solo se tienen en cuenta las notas de las preguntas de selección múltiple, las cuales las genera la plataforma automáticamente.

## IV. SELENEVIS

SelenVis es una aplicación Web que trabaja con los registros de datos de la Plataforma Selen Unicauca, una plataforma para la disposición de cursos en línea de la Universidad del Cauca. La Aplicación Web presenta un mecanismo de visualización de datos, para la identificación y seguimiento gráfico de los comportamientos con sospecha de fraude de los estudiantes con los diferentes Cursos. Está

pensada para brindar apoyo a los docentes encargados de los diferentes cursos, para que puedan brindar la asesoría y el acompañamiento a los estudiantes, especialmente a los que presentan comportamientos denominados sospechosos. Esto permitiéndoles una interfaz del mecanismo de visualización propuesto, que se compone de 4 vistas principales: “login”, “Sospechosos”, “Estadísticas”, “Comportamientos” y una vista de administración de docentes y/o administradores de la aplicación. A continuación, se describe cada una de las vistas:

### A. Vista “Login”

Al ingresar a la aplicación web la primera vista en mostrarse es la vista de Inicio de sesión (login), en la cual se requiere del correo electrónico y la contraseña del usuario para poder ingresar a las funcionalidades de la misma. Los usuarios solo pueden ser registrados a través del módulo de administración descrito en la sección IV-E. La figura 1 muestra la interfaz web de esta vista.

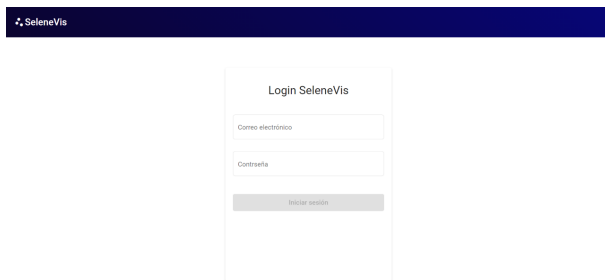


Fig. 1. Vista de login

### B. Vista “Sospechosos”

El objetivo de esta vista presentada en la figura 2 es proporcionar al docente una representación gráfica mediante dendrogramas<sup>1</sup> en los cuales se representa la similitud entre grupos de estudiantes de acuerdo a las respuestas y diferencia de tiempos que obtuvieron durante el curso o examen que se selecciona en la lista de opciones.

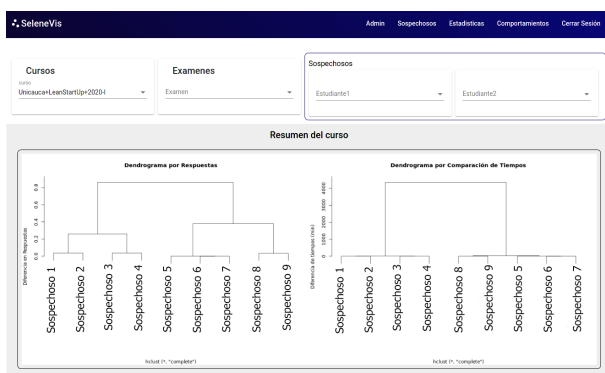


Fig. 2. Vista sospechosos - resumen general

<sup>1</sup>Representación gráfica de datos en forma de árbol, se divide en subcategorías hasta llegar a la información final.

En la parte inferior izquierda se encuentra un dendrograma que representa la cercanía de cada estudiante con respecto a las respuestas enviadas. En la aplicación agrupa a los estudiante más cercanos. Por ejemplo los Sospechosos 5, 6 y 7 han sido agrupados porque tiene una cercanía de 0. Es decir tienen un 100% de de coincidencia en sus respuestas, tanto correctas como incorrectas. De la misma manera, en la parte inferior derecha se muestra un dendrograma, pero realizado teniendo en cuenta la cercanía de los estudiantes en el envío de sus respuestas. Nuevamente, los sospechosos 5, 6 y 7 han sido agrupados debido a que enviaron sus exámenes muy cerca en el tiempo. La generación de las gráficas de dendrogramas se realizaron con base en el trabajo previo presentado en [21], [22].

Otra característica de la pestaña de “Sospechosos” es la comparación de dos estudiantes. Esta vista permite obtener directamente la similitud entre estudiantes de acuerdo su comportamiento en el momento de presentar un examen en la plataforma. La figura 3 muestra que al seleccionar el estudiante sospechoso 5 y el sospechoso 6, es posible observar el resultado de su comparación, aquí se observa la fecha de presentación del examen, el número de coincidencias en respuestas y la diferencia de tiempo de envío del examen.

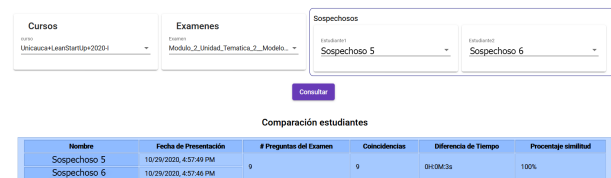


Fig. 3. Similitud estudiantes

### C. Vista “Estadísticas”

Esta vista permite ver las estadísticas generales del curso de acuerdo a las interacciones totales que los estudiantes han tenido dentro del mismo. Las interacciones que se han tenido en cuenta son: interacción con los contenidos, vídeos, foros y exámenes. También se ha tenido en cuenta el número de sesiones que el estudiante presenta en todo el curso. En figura 4 se muestra un pantallazo de dicha pestaña. En la parte izquierda se muestra un diagrama de barras con la cantidad de vídeos, contenidos, foros, exámenes y sesiones que se han presentado en todo el curso. En la parte derecha se muestra una tabla con el listado de estudiantes que pertenecen al curso junto con sus interacciones particulares. Esta tabla puede ser ordenada alfabéticamente por nombre de estudiante o puede ser ordenada de mayor o menor por cada una de las interacciones mencionada anteriormente.

Adicionalmente en la parte inferior, se observa una gráfica con una línea de tiempo en la cual se muestran el total de interacciones durante el curso. Esta gráfica presenta la cantidad de interacciones de los estudiante día a día. Cabe destacar que esto indica cuál es el comportamiento que tienen los estudiantes durante el curso. Como se puede observar en la parte inferior de la figura 4. En el ejemplo de la figura 4, en el historial de interacciones, se puede observar con

círculos de color verde los días en que hubo exámenes y se aprecia picos de interacciones un día antes o el mismo día del examen.

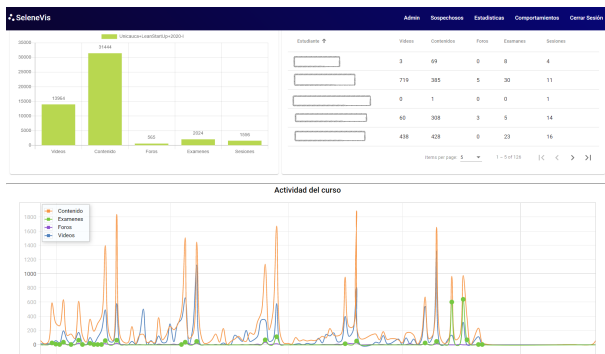


Fig. 4. Vista estadísticas

También dentro de esta vista es posible consultar las estadísticas de un estudiante en particular. El resultado de la consulta se muestra en un diagrama de barras. Solo basta seleccionar un estudiante en la lista y se puede apreciar el resultado en pantalla. La figura 5 muestra el resultado de consultar un estudiante.

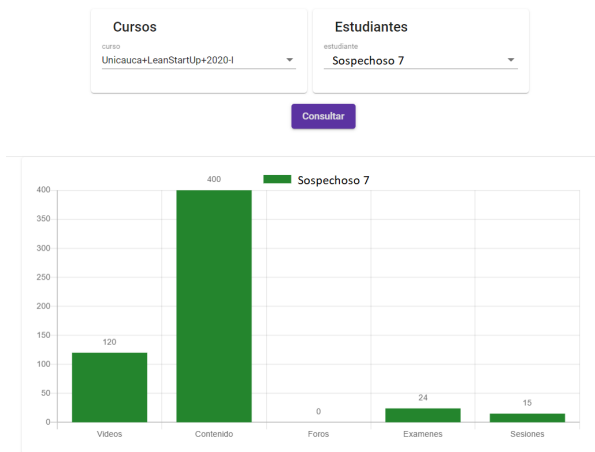


Fig. 5. estadísticas estudiante

#### D. Vista “Comportamientos”

En esta vista es posible consultar y comparar los comportamientos de los estudiantes en el curso. La figura 6 muestra los dos tipos de gráficas que se encuentran en esta vista. En la parte superior están las estadísticas del estudiante consultado y en la parte inferior se muestra una línea de tiempo orientada a grafos, donde se observa la relación entre los diferentes nodos(actividades), la cual es un resumen de las actividades que el estudiante realizó dentro del curso y en la cual se describe una secuencia ordenada de los pasos que hizo el estudiante bien sea por día o por sesión.

Una vez consultado el estudiante, encima de las gráficas aparece una la opción para escoger la forma en como se desea mostrar los resultados. Es decir, si se escoge la categoría de “día” aparecerá la lista con opciones de los

días en que el estudiante interactuó con el curso y en caso de escoger “sesion” aparece la lista de sesiones que tiene el estudiante. Una vez escogida la opción se despliega la información estadística y su comportamiento representado por grafos. Para la gráfico de representación de comportamientos por grafos nos basamos en el trabajo [], donde se propone visualizar los comportamientos de los estudiantes mediante grafos. Nosotros tuvimos en cuenta su propuesta y la implementamos en la aplicación que realiza las gráficas de manera automática y en tiempo real, algo que no se encuentra registrado en la literatura.



Fig. 6. Vista comportamientos

#### E. Vista “Admin”

La vista “Admin” (administrador), presenta una interfaz que permite listar usuarios registrados y la gestión de nuevos usuarios. Para ello se tiene un formulario de registro y asignación de roles, estos roles son de docente y administrador. Para las respectivas opciones de registro, edición, eliminación se tiene un formulario que lista los usuarios registrados y además añade los botones para la gestión. Este procedimiento es obligatorio para el posterior inicio de sesión de ingreso a la aplicación. El administrador es quien crea las cuentas de los docentes y asigna los cursos según se encuentren asignados en la plataforma Selene. De esta manera los docentes tienen acceso sólo a los cursos en los cuales estén a cargo.

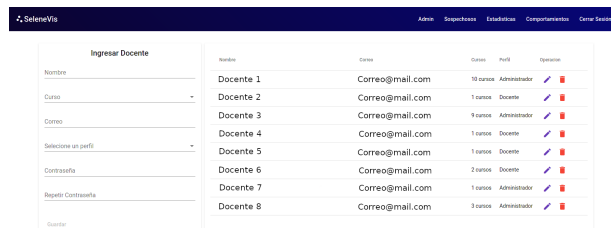


Fig. 7. Vista de Administrador

## V. CONCLUSIÓN Y TRABAJOS FUTUROS

En este artículo demostramos una herramienta para la visualización de datos relacionados a los comportamientos de los estudiantes dentro de una plataforma MOOC "Selene Unicauca".

La manera en que los MOOC se están incorporando en el ámbito universitario es a través de los MPOC y SPOC que son cursos que al ser privados permiten tener un mejor control sobre los estudiantes.

Los cursos con reconocimiento académico son más susceptibles a que los estudiantes cometan fraude en sus exámenes.

Los MOOC al tener reconocimiento académico hace que los docentes necesiten tener un mayor control y seguimiento de sus estudiantes, lo cual no es posible en las plataformas actuales destinadas para MOOC, ya que este tipo de cursos están destinado al auto aprendizaje.

El desarrollo de trabajos como "Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics" desarrollado por el estudiante de Doctorado Daniel Jaramillo-Morillo y el doctor José Ruipérez-Valiente [1], permiten un adelanto a la identificación de estudiantes sospechosos de fraude en el entorno de los SPOCs a partir de métricas de similitud de respuestas y tiempos de envío. Para ello se ha retomado su trabajo y se ha generado una conexión que permite apreciar sus resultados actualizados en el mecanismo de visualización que se ha desarrollado.

Para la implementación del mecanismo de Visualización "SELENEVIS", se ha retomado los trabajos de "Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics" desarrollado por el estudiante de Doctorado Daniel Jaramillo-Morillo y el doctor José Ruipérez-Valiente [1] y un trabajo de maestría de la Universidad del Cauca. Para el primer caso se ha realizado una conexión que permite incorporar los resultados en el mecanismo de visualización, haciéndolo más accesible y permitiendo un análisis complementario a partir de datos estadísticos a lo largo del curso. En el segundo Caso, se ha retomado una idea de seguimiento de las interacciones de los estudiantes mediante conexiones de Nodos. Para lo cual se ha realizado la implementación gráfica de la estrategia en el mecanismo de visualización, esto a partir de la aplicación Web que permite reunir más información y generar un esquema de grafos interconectados que visualiza el seguimiento en línea cronológica de las interacciones de los estudiantes con la plataforma y el material del curso, en las diferentes sesiones o un día específico a lo largo del avance del curso.

El mecanismo de visualización "SELENEVIS", es una aplicación web, que proporciona a los docentes y administradores una herramienta para dar seguimiento del rendimiento de los diferentes cursos que tenga asignados, permite seguir el comportamiento de los estudiantes respecto a sus interacciones con la plataforma, y su comportamiento a través de las diferentes sesiones o un día específico.

Agrupar diferentes vistas que permiten diferentes niveles de abstracción. Estas son: La vista de administrador, la vista de sospechosos, la vista de comparación y la vista de Estadísticas.

Para la implementación del mecanismo, los datos de los estudiantes se obtienen del archivo tracking.log que ofrece la plataforma SELENE UNICAUCA que es una instancia de la OpenEdx implementada por la Universidad del Cauca. Una vez estudiando los campos de este archivo, se retoma el trabajo propuesto por el estudiante de doctorado Daniel Jaramillo Morillo en su tesis de maestría [22]. A partir de esta información se realiza una categorización de estos datos, que permiten adaptar el modelo de datos que permite la construcción de los diferentes componentes que finalmente se integran en el mecanismo de visualización. Permitiendo consultar la información en tiempo Real (Siendo el tiempo real, consultar los datos actualizados del día inmediatamente anterior al día en el que se hace las respectivas consultas) desde la aplicación web del mecanismo de visualización.

Mediante el mecanismo de visualización se puede realizar un seguimiento de las estadísticas generales a nivel del curso o a nivel de cada uno de los estudiantes. También es posible administrar los usuarios docentes y administradores, identificar estudiantes con sospecha de fraude a partir de las métricas de similitud de respuestas y similitud de tiempos de envío de los cuestionarios.

Con esta información es posible hacer una revisión específica de los estudiantes detectados como sospechosos a partir de una vista de comparación que muestra las estadísticas de los estudiantes acorde a la categorización de los datos en Contenidos, vídeos, Foros, Exámenes y sesiones. Además, esta vista de comparación permite hacer seguimiento de las interacciones de los estudiantes a lo largo de los diferentes días transcurridos del curso o en función de las sesiones. Esto mediante un esquema de grafos conectados que permite seguir el orden en el que interactuó el estudiante ya sea en el día o por sesión.

Se recomienda en trabajos futuros considerar la nota del examen final y si el estudiante aprobó o no el curso, con el fin de determinar, por ejemplo, si los casos de free-riding ganan o no el examen.

Finalmente se recomienda considerar las notas de preguntas que no sean de selección múltiple o la inclusión de las mismas, y analizar complementariamente que implicaciones tienen sobre el rendimiento de los estudiantes y sus comportamientos con sospecha de fraude, estos sobre diferentes casos de estudio, ya que este tipo de notas no se tuvieron en cuenta.

## REFERENCES

- [1] D. Jaramillo-Morillo, J. Ruipérez-Valiente, M. F. Sarasty, and G. Ramírez-Gonzalez, "Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics," vol. 17, no. 1, p. 45.
- [2] D. Jaramillo, M. Sarasty, G. Ramirez-Gonzalez, and M. Pérez-Sanagustín, "Follow-up of learning activities in open edx: A case study at the university of cauca," *link.springer.com*, pp. 217–222, 05 2017.
- [3] V. Mailhes and J. Raspa, "Mooc: De la revolución educativa a la supervivencia," p. 17.

- [4] F. Dernoncourt, C. Taylor, K. Veeramachaneni, S. Wu, S. Halawa, and U.-M. O'Reilly, "Moocviz: A large scale, open access, collaborative, data analytics platform for moocs," *www.researchgate.net*, 11 2013.
- [5] M. Santamaría Lancho, "Moocs y spocs (small private online courses): sus posibilidades para la formación del profesorado," vol. 1, no. 1, p. 6.
- [6] D.-S. Chen and G.-P. Yan, "MPOC experiential learning model of engineering education based on distributed flip in local engineering colleges and universities in "post MOOC" era," vol. 20, no. 6, pp. 1375–1379.
- [7] C. Shi, S. Fu, Q. Chen, and H. Qu, "VisMOOC: Visualizing Video Clickstream Data from Massive Open Online Courses," p. 2.
- [8] J.-S. Wong and X. L. Zhang, "MessageLens : A visual analytics system to support multifaceted exploration of mooc forum discussions," *Visual Informatics*, vol. 2, pp. 37–49, Mar. 2018.
- [9] X. Li, C. Men, F. Zhang, and Z. Du, "A Smart Visual Analysis Solution for MOOC Data," in *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, (Orlando, FL), pp. 101–106, IEEE, Nov. 2017.
- [10] K. Sarajeva, "Ict for anti-corruption , democracy and education in east africa," in *Democracy and Education in East Africa*, 2013.
- [11] W. Jobe, "No university credit, no problem? exploring recognition of non-formal learning," in *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, pp. 1–7, Oct 2014.
- [12] P. McGee, "Supporting Academic Honesty in Online Courses," *The Journal of Educators Online*, vol. 10, Jan. 2013.
- [13] B. Zirger, E. Rutz, D. Boyd, J. Tappel, and V. Subbian, "Creating pathways to higher education: A cross-disciplinary mooc with graduate credit," pp. 1–5, 03 2014.
- [14] C. Gherib, J.-M. Dujardin, and D. Verpoorten, "MOOCS IN BUSINESS ADMINISTRATION – AN OVERVIEW OF ASSESSMENT PRACTICE," p. 15.
- [15] T. Staubitz, R. Teusner, J. Renz, and C. Meinel, "An Experiment in Automated Proctoring," p. 12, 2016.
- [16] J. Miguel Moneo, S. Caballé, and J. Prieto-Blazquez, "Providing information security to mooc: Towards effective student authentication," pp. 289–292, 09 2013.
- [17] R. Woolsey, "Student verification system for online assessments: Bolstering quality and integrity of distance learning," *Journal of Industrial Technology*, vol. Volume 27, 09 2011.
- [18] Y. Atoum, L. Chen, A. X. Liu, S. D. H. Hsu, and X. Liu, "Automated online exam proctoring," *IEEE Transactions on Multimedia*, vol. 19, pp. 1609–1624, July 2017.
- [19] S. Prathish, A. N. S., and K. Bijlani, "An intelligent system for online exam monitoring," in *2016 International Conference on Information Science (ICIS)*, pp. 138–143, Aug 2016.
- [20] M. Cote, F. Jean, A. B. Albu, and D. Capson, "Video summarization for remote invigilation of online exams," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–9, March 2016.
- [21] D. Jaramillo-Morillo, M. Solarte, and G. R. González, "Estrategia de seguimiento a las actividades de aprendizaje de los estudiantes en cursos en línea masivos y privados (MPOC) con reconocimiento académico en la universidad del cauca," p. 20.
- [22] D. Jaramillo-Morillo, M. S. Sarasty, and G. Ram, "Follow-up of learning activities in open edX: A case study at the university of cauca," p. 6.