



## Analítica de aprendizaje como estrategia de apoyo al aula invertida en cursos de programación: una revisión sistemática de literatura

The learning analytics to support of the flipped classroom in programming courses: a systematic literature review

José Miguel Llanos Mosquera



Víctor Andrés Bucheli Guerrero



Universidad del Valle, Colombia

OPEN ACCESS

Recibido: 01/12/2020

Aceptado: 17/02/2021

Publicado: 16/04/2021

Correspondencia de autores:

jose.llanos@correounivalle.edu.co



Copyright 2020  
by Investigación e  
Innovación en Ingenierías

### Resumen

**Objetivo:** Identificar los métodos y las técnicas que se están utilizando en la analítica de aprendizaje, como apoyo al aula invertida para los cursos de programación. **Metodología:** Se realizó la búsqueda de artículos publicados en revistas y conferencias científicas, a partir de herramientas bibliográficas que ayudaron a construir el corpus de referencias. Luego se analizó la información más significativa de los documentos, para el desarrollo de la revisión sistemática. **Resultados:** Esta revisión identificó que los métodos y las técnicas han contribuido en la preparación de los estudiantes, la definición de estrategias para el aprendizaje de la programación, herramientas de detección de errores en el código fuente y sistemas de tutorías inteligentes. Mejorando la retención de los estudiantes, los hábitos de estudio y las experiencias de aprendizaje. **Conclusiones:** Las investigaciones han analizado los datos almacenados en sistemas de gestión de aprendizaje (LMS), con el fin de generar modelos de predicción que permitan mejorar la formación en línea, el rendimiento académico y la retroalimentación personalizada. Todos estos aportes han apoyado a las áreas de conocimiento de aprendizaje automático y minería de datos, con la solución de problemas relacionados al aprendizaje supervisado, no supervisado y la minería de datos educativa.

**Palabras clave:** Analítica de aprendizaje, aula invertida, cursos de programación, aprendizaje automático, minería de datos, minería de datos educativa.

### Abstract

**Objective:** Identify the methods and techniques used in learning analytics to support the flipped classroom in the programming courses. **Methodology:** The systematic search was carried out of articles published papers in journals and conferences scientific from bibliographies tools than helped to build the corpus of references. Then analyzed the information more significant of the documents for the development of the systematic review. **Results:** This review identified the methods and the techniques that contribute to the preparation of the student, the definition of strategies for learning programming, tools for the detection of errors in the source code, and system of intelligence tutorials. Improving student retention, the habits of study, and the experiences of learning. **Conclusions:** The investigations have analyzed the stored data in the learning management systems (LM) to generate predictive models to improve online training, academic performance, and personalized feedback. All these contributions have supported the machine learning knowledge areas and data mining to solve problems related to supervised learning, unsupervised learning, and educational data mining.

**Keywords:** Learning analytics, flipped classroom, programming courses, machine learning, data mining, educational data mining.

**Como citar (IEEE):** J. Llanos-Mosquera, y V. Bucheli-guerrero. "Analítica de aprendizaje como estrategia de apoyo al aula invertida en cursos de programación: una revisión sistemática de literatura", *Investigación e Innovación en Ingenierías* vol. 9, nº 1, 114-135, 2021. DOI: <https://doi.org/10.17081/invinno.9.1.4464>

## Introducción

A inicios del siglo XXI la educación superior dio un giro importante a los procesos de enseñanza-aprendizaje, porque incorporó herramientas web que permitieron mejorar la participación de docentes y estudiantes. En los últimos años muchas universidades han incorporado entornos virtuales de aprendizaje y sistemas de gestión de aprendizaje (LMS por sus siglas en inglés) para desarrollar sus cursos a través de estas herramientas, según el Online Business School se espera que a partir del 2019 el 50% de las clases de educación superior en el mundo, se impartan a través de esta modalidad de aprendizaje [1]. Sin embargo, todas estas tecnologías generan huellas o senderos digitales (registros, acciones y patrones) que se deben analizar, con el fin de mejorar la enseñanza, el aprendizaje y el éxito del estudiante [2].

Teniendo en cuenta los desafíos relacionados a la huella digital, apoyo al estudiante y rendición de cuentas en la educación superior, apareció la analítica de aprendizaje (LA del inglés Learning Analytics), un área de interés que permitió encontrar solución a todos estos retos [3]. LA utiliza los conocimientos obtenidos durante el análisis de los datos, para mejorar el proceso de aprendizaje y apoyar el entorno en el que se desarrolla, de esta forma se puede monitorear las actividades de los estudiantes y pronosticar el desempeño temprano, mejorando la retención y retroalimentación de los aprendices [4]. Actualmente está llamando la atención de académicos e investigadores, porque existe la necesidad de comprender y mejorar la enseñanza, el aprendizaje, la personalización y la adaptación. Aunque se encuentra en las primeras etapas de investigación e implementación, ya existe una organización denominada Society for Learning Analytics Research, que surgió para fomentar una comunidad de investigación y contribuir desde la analítica de aprendizaje a la educación [2].

La analítica de aprendizaje y la minería de datos educativa (EDM del inglés Educational Data Mining) son dos conceptos que están estrechamente relacionados, LA utiliza varios métodos que incluyen técnicas para predecir, agrupar, descubrir y separar datos; también aplica modelos que permiten abordar problemas que afectan el aprendizaje de los estudiantes. Mientras que EDM es un subconjunto de la minería de datos, se enfoca en desarrollar e implementar métodos de análisis, con el objetivo de examinar patrones relacionados con las acciones de los estudiantes y su entorno educativo, asimismo utiliza algoritmos de minería de datos que permiten resolver problemas educativos generando respuesta sistemáticas y automatizadas para los aprendices y los ambientes de formación [5].

El aula invertida (FC del inglés Flipped Classroom) es un método pedagógico, combina teorías de aprendizaje que alguna vez fueron consideradas incompatibles, como el aprendizaje basado en problemas y conferencias derivadas en métodos de instrucción [6]. También se define como una técnica educativa con dos partes, la primera parte incluye actividades interactivas de aprendizaje grupal, que se realizan dentro del aula de formación a través de tareas y resolución de problemas; la segunda parte define instrucciones individuales que se desarrollan en un computador fuera del aula de clase a través de videoconferencias asíncronas [6]. El enfoque principal de esta estrategia se centra en hacer la formación al revés, esto implica que el docente debe impartir conferencias antes de la clase, en forma de videos pregrabados y el tiempo de formación lo utiliza con sus estudiantes en actividades que involucran aprendizaje activo, colaboración, interacción y resolución de problemas [7].

Aunque la analítica de aprendizaje y el aula invertida son temas de interés, gran parte de la información existente se encuentra en blog, periódicos y revistas en línea en lugar de artículos académicos y conferencias [7]. No obstante, existen pocos trabajos investigativos que permiten evaluar de forma objetiva los resultados de aprendizaje de los estudiantes, para guiar las actividades de enseñanza a través de LA y

FC [6]. Todo esto ha motivado a muchos científicos a desarrollar herramientas para profesores y estudiantes que permitan medir sus efectos, desafíos y efectividad, con el fin de proponer las mejores prácticas [3, 6]. Sin embargo, hasta el momento no es claro cuáles métodos, técnicas y herramientas basadas en la analítica de aprendizaje, han apoyado al aula invertida, con el fin de mejorar los procesos de enseñanza en los cursos de programación, por este motivo se propone el desarrollo de esta revisión sistemática de literatura.

El artículo se compone de cuatro secciones, en primer lugar, se presenta el estado del arte que contiene los métodos y técnicas de LA como apoyo al FC; luego se define la metodología utilizada en la revisión que incluye: las preguntas de investigación, fuentes de información, cadenas de búsqueda, criterios de inclusión y exclusión y el análisis del corpus de referencias; posteriormente se presentan los resultados y discusión; finalmente se encuentran las conclusiones.

## Estado del Arte

En esta sección se responde la pregunta: ¿Qué métodos y técnicas de analítica de aprendizaje sirven como apoyo al aula invertida?, teniendo en cuenta el corpus de referencias se procedió a realizar el análisis de los documentos, identificando métodos y técnicas de LA que sirven de apoyo al FC. Cada uno de estos elementos se organizaron en las categorías de: clasificación, regresión, agrupación y minería de datos para la comprensión de sus aportes.

## Clasificación

La clasificación se utiliza para generar resultados a partir de etiquetas, generando respuesta al problema con un conjunto finito de posibles resultados, durante el análisis de la información se encontraron los siguientes métodos y técnicas.

## Métodos

Los métodos más utilizados de clasificación que se identificaron en el análisis de los artículos del corpus de referencias son: Naive Bayes, árboles de decisión, máquinas de vectores de soporte, proceso Gaussiano y vecinos más cercanos. A continuación, se describen los principales aportes que se han generado en estos métodos.

## Naive Bayes

En este clasificador se encontraron diversos aportes relacionados a estudiantes en riesgo, asesoramiento electrónico y competencias de liderazgo. Se analizó una investigación que define un modelo de predicción basado en Naive Bayes, para identificar los estudiantes en riesgo, a partir de los datos recolectados en un LMS del proceso de formación y las evaluaciones [8]. Otra investigación propone un software inteligente de asesoramiento electrónico, utiliza este método para identificar y extraer palabras claves del texto, que contienen emociones como alegría, sorpresa y recuerdo [9]. Finalmente se identificó una investigación que predice competencias de liderazgo, utilizando las redes sociales como herramienta de colaboración y Naive Bayes para identificar el flujo de las interacciones sociales [10].

## Árboles de decisión (decision trees)

Con respecto a los árboles de decisión se han desarrollado varias investigaciones que incluyen la participación, preparación y evaluación del estudiante. En un proyecto se desarrolló un modelo basado en árboles de decisión, con el fin identificar factores de participación del estudiante, a través de actividades desarrolladas en LMS, que incluyen colaboración, visualización y discusión [9]. Otro proyecto analiza la relación entre la preparación de los estudiantes antes de clase y los puntajes obtenidos en los exámenes. Incluye un árbol de decisión, que permite clasificar los aprendices a partir de los registros y puntajes obtenidos en las pruebas [11]. Asimismo, se analizaron varios artículos que predicen el rendimiento académico y tendencias de aprendizaje de los estudiantes, a partir de los datos recopilados durante el proceso de formación. Definen modelos que incluyen árboles de decisión para predecir los estudiantes que se encuentran en riesgo, con base a variables demográficas, resultados de exámenes, calificación final del curso y retroalimentación del docente [12, 13, 14].

De igual forma se analizaron otras investigaciones que utilizan árboles de decisión, con el fin de investigar el comportamiento de aprendizaje en línea de los estudiantes [8] o para identificar la representación del conocimiento, a partir de recursos como tesauros, mapas conceptuales y semantización, que hacen parte del proceso formativo [15]. También existen documentos donde se incluyen modelos con árboles de decisión que permiten detectar de forma automática las emociones generadas en el texto por los estudiantes, apoyando la toma de decisiones para las asesorías [16]. Asimismo se analizaron dos trabajos investigativos que contribuyen a la enseñanza de la programación, el primero define un método que incluye reglas para corregir composiciones sintácticas, a partir de los errores cometidos por el estudiante [17] y el segundo presenta un método automático basado en árboles de decisión para analizar los mensajes enviados por los estudiantes en los debates, su profundidad y similitud con otros mensajes [18].

## Máquinas de vectores de soporte (SVM)

Las máquinas de vectores de soporte son otro método importante de clasificación, el análisis se centró en la predicción del rendimiento académico, el abandono y el riesgo de los estudiantes. Se identificaron varios trabajos que utilizan SVM, uno examina los clics generados en el material de estudio, con el fin de identificar el número de visitas al contenido y la cantidad de veces que se observa un video [19]; otro predice el rendimiento académico de los estudiantes, a partir de los datos recopilados en LMS y las actividades relacionadas con los resultados de aprendizaje [8] y un estudio predice los estudiantes en riesgo, identificando patrones de aprendizaje en el desarrollo de cuestionarios y la retroalimentación del docente [16]. También se analizaron otras investigaciones que utilizan SVM para predecir la calidad de la colaboración entre grupos de trabajo [20]; el control de la reproducción de videos a partir de los gestos generados por los estudiantes en sus teléfonos móviles [21]; la toma de decisiones para generar asesoría con base a los antecedentes emocionales de los aprendices [9] y el análisis automático del contenido enviado en los mensajes por los estudiantes [13].

## Proceso Gaussiano (gaussian processes)

El proceso gaussiano es otro método utilizado para clasificación, los aportes más significativos encontrados durante el análisis, se centran en la identificación de estudiantes en riesgo y conversaciones en chat. En el primer tema se analizó una investigación que identifica a través de gaussian processes los estudiantes en riesgo, para generar la intervención oportuna de acuerdo con los resultados obtenidos en las evaluaciones [8]. Con respecto a las conversaciones en chat, se analizó un artículo que utiliza el proceso Gaussiano y técnicas de procesamiento del lenguaje natural, para identificar de forma automática los vínculos y contribuciones generadas a partir de los hilos de discusión [22].

## Vecinos más cercanos (nearest neighbors)

En la clasificación también se utilizan los vecinos más cercanos, que han contribuido en el desarrollo de software y el análisis de contenido. Se desarrolló un sistema inteligente que utiliza vecinos más cercanos, para la extracción de características y el análisis del contenido consumido por los estudiantes a través de su teléfono móvil, que se encuentra disponible en cursos MOOC [21]. Otra investigación utiliza nearest neighbors para analizar el contenido generado por los estudiantes en debates creados en LMS [13].

## Técnicas

Las técnicas más utilizadas para clasificación que se identificaron en el análisis de los artículos del corpus de referencias son: validación cruzada y aumento del gradiente extremo. A continuación, se describen los principales aportes que se han generado para estas técnicas.

## Validación Cruzada (cross-validation)

En la técnica de validación cruzada se identificaron varios aportes, por ejemplo, una investigación utiliza la validación cruzada para analizar los indicadores y el comportamiento generado por el estudiante, en las pruebas unitarias que se aplican al código fuente de los talleres de programación [23]. Otro trabajo presenta un sistema de tutoría inteligente con validación cruzada, que se encarga de detectar los errores más comunes de codificación generados por los estudiantes y proporciona los respectivos comentarios [14]. Asimismo, existen artículos investigativos que utilizan esta misma técnica, con el fin de predecir el rendimiento académico e identificar los estudiantes en riesgo, con base a los resultados obtenidos en las evaluaciones [8] o para analizar el desempeño de los aprendices y mejorar su retención, a partir de las estrategias de aprendizaje utilizadas y los patrones generados en las actividades desarrolladas en el curso [24]. También existen estudios de caso que investigan la relación entre la preparación de los estudiantes y los resultados obtenidos en los exámenes, emplean la validación cruzada para predecir los puntajes y el tiempo empleado para resolver los cuestionarios [15].

Un estudio utiliza la validación cruzada con el fin de clasificar y evaluar si un aprendiz es líder o no, a partir de un conjunto de características que incluyen comportamientos y actividades generadas en redes sociales [10]. Otro artículo utiliza esta técnica para analizar si los niveles de colaboración y la cantidad de interacciones pueden afectar el rendimiento académico de los estudiantes [19]. De la misma forma,

existe una investigación que utiliza la validación cruzada para analizar de forma automática el contenido generado por los aprendices, en los foros de discusión [13]. Finalmente, se utiliza la técnica para detectar y analizar las características generadas en los chat por los participantes, a partir de las palabras [22].

### **Aumento del gradiente extremo (extreme gradient boosting)**

En esta técnica de clasificación se analizó una investigación que propone mejorar las experiencias de aprendizaje de los estudiantes, implementa la técnica de aumento del gradiente extremo para predecir las calificaciones de los aprendices, con el fin de generar la intervención oportuna y disminuir la deserción [8].

### **Regresión**

La regresión se encarga de predecir elementos continuos, con base a la función de entrada del modelo, en esta categoría se identificaron el método de mínimos cuadrados y las técnicas de regresión lineal y regresión logística.

### **Métodos**

Los métodos más utilizados de clasificación que se identificaron en el análisis de los artículos del corpus de referencias son: mínimos cuadrados, árboles de decisión, máquinas de vectores de soporte, proceso gaussiano y vecinos más cercanos. A continuación, se describen los principales aportes que se han generado para estos métodos.

### **Mínimos Cuadrados (least squares)**

Existe una investigación que utiliza mínimos cuadrados para predecir cómo la participación y el comportamiento en línea, pueden afectar los resultados del proceso de formación en cursos de programación [25]. Otros estudios utilizan el método y el aprendizaje autorregulado, con el propósito de explorar la relación entre el entorno de aprendizaje en línea y el rendimiento académico [26] o para descubrir factores que influyen en la autorregulación [27]. Asimismo, se encontró un artículo que utiliza un modelo con mínimos cuadrados con el fin de evaluar la confiabilidad y validez de las percepciones generadas por los maestros, en el uso de herramientas que ayudan al aprendizaje de la programación [28]. También se encontró un documento que implementa el método para analizar cómo las condiciones de las herramientas Moodle influyen en el éxito académico [29]. Finalmente, un estudio utiliza mínimos cuadrados para analizar la correlación existente entre la facilidad de uso y comportamientos de uso de la herramienta Moodle por parte de los estudiantes [30]. Con respecto a los métodos de árboles de decisión, máquinas de vectores de soporte, proceso gaussiano y vecinos más cercanos, es importante aclarar que los aportes generados por los artículos del corpus de referencias, se encuentran descritos en los respectivos párrafos de la categoría de clasificación.

## Técnicas

Las técnicas más utilizadas para regresión que se identificaron en el análisis de los artículos del corpus de referencias son: regresión lineal y regresión logística. A continuación, se describen los principales aportes que se han generado para estas técnicas.

### Regresión Lineal (linear regression)

En esta técnica se analizaron varias investigaciones, una de ellas utilizó regresión línea para analizar los efectos de cambiar la evaluación del aula tradicional al aula invertida en un curso de ingeniería de software [31]. Otra investigación combinó la regresión, variables demográficas y el envío de proyectos, para identificar y predecir los estudiantes que se encuentran en riesgo [8]. Asimismo, se consideró una investigación que utiliza un modelo de regresión lineal para examinar la medida en que las condiciones de los LMS pueden influir en el éxito académico de los estudiantes, se analizó el acceso a los foros y las tareas desarrolladas por los estudiantes [29]. Varios documentos utilizan esta técnica para estudiar la relación entre los comportamientos de aprendizaje en línea y los resultados de enseñanza [32], por ejemplo, existe un estudio de caso donde se explora dicho comportamiento a través de un tablero digital que incluye videoclips y puntajes obtenidos en los exámenes [33]; otro artículo realiza el mismo análisis a través de estrategias autoinformadas, que permiten identificar patrones en las actividades desarrolladas por los estudiantes y sus experiencias de aprendizaje [24].

Se analizaron otras investigaciones que utilizan la regresión lineal con el fin de descubrir si la cantidad de preguntas en las evaluaciones pueden influenciar en el comportamiento y rendimiento generar del aprendiz [34]; de la misma forma se utilizó los registros almacenados en cursos MOOC, para analizar la correlación entre las tareas obligatorias y opcionales que debe desarrollar el estudiante [35]; en último lugar, se encontró un trabajo que analiza la relación entre las anotaciones de video, los enfoques de aprendizaje de los estudiantes y el rendimiento académico, verificando la relación entre las calificaciones y el nivel de interacción con las anotaciones de videos [36].

### Regresión Logística (logistic regression)

Se analizó una investigación que utiliza la regresión logística para predecir si el estudiante abandonará un curso o recibirá un certificado de finalización, las variables utilizadas para la predicción son personalidad y colaboración [19]; existe un estudio que utiliza un modelo con regresión logística que permite identificar patrones de éxito de los aprendices a través de calificaciones y las experiencias de aprendizaje [24]; otras investigaciones utilizan regresión logística para identificar la relación entre los elementos de los LMS y el rendimiento académico, a partir de las características de los estudiantes y datos derivados del seguimiento [8, 29]. Existe una investigación que utiliza la regresión logística para identificar patrones de aprendizaje, utilizados por los estudiantes en cursos online, en el estudio se analizan las visitas por días, los cuestionarios y la retroalimentación del docente [16]. También se utiliza esta técnica para priorizar las necesidades de apoyo de los estudiantes de programación y contribuir al docente en la toma de decisión del proceso formativo [37]. Los últimos temas analizados, hacen referencia a sistemas inteligentes de aprendizaje y análisis de competencias de liderazgo. En el primer tema se encontró AttentiveLearner, un sistema inteligente de aprendizaje móvil que utiliza la regresión para predecir los eventos de “vagancia”

de los estudiantes en las sesiones MOOC y el desempeño en los cuestionarios [21]; en el segundo tema se utilizó la misma técnica para identificar evidencias de liderazgo a un conjunto de datos, que contiene actividades realizadas en redes sociales [10].

## **Agrupación (clustering)**

La agrupación genera modelos o patrones a partir de datos recolectados, permitiendo que se pueda mejorar el objetivo con base a sus funciones. Es importante aclarar que, durante el análisis de la información solo se encontraron métodos para agrupación.

## **Métodos**

Los métodos más utilizados para agrupación que se identificaron en el análisis de los artículos del corpus de referencias son: agrupamiento jerárquico y método de Ward. A continuación, se describen los principales aportes que se han generado para estos métodos.

## **Agrupamiento jerárquico (hierarchical clustering)**

Un trabajo investigativo empleó el agrupamiento jerárquico aglomerativo para detectar patrones en el comportamiento y estrategias de aprendizaje adoptadas por el estudiante, con la agrupación de los estudiantes y las secuencias de aprendizaje [38]. Asimismo, una investigación utilizó el mismo método para identificar el comportamiento del estudiante a partir de los datos generados en los recursos en línea, las sesiones de estudio y las estrategias de formación [39]. Otra investigación utilizó el agrupamiento para identificar patrones de secuencias en actividades desarrolladas por los estudiantes, que estaban relacionadas a fijación de objetivos, planificación de estrategias y búsqueda de ayuda [24]. También se analizó un documento que empleó el agrupamiento jerárquico, para detectar las tácticas y estrategias de gestión de tiempo utilizadas por los estudiantes para el desarrollo de sus tareas de formación [40]. Existen estudios que utilizan la agrupación jerárquica para descubrir las trayectorias de rendimiento de los estudiantes a partir de la autoeficacia, comportamientos de autorregulación y comprensión de contenidos [41]; o para evaluar el comportamiento de los estudiantes, con el fin de identificar las diferentes habilidades que se relacionan con el rendimiento académico [42]; también para analizar la participación de los estudiantes, con la ayuda del procesamiento del lenguaje natural y la agrupación jerárquica, que se encargan de extraer las interacciones y colaboración de los participantes, a través de los mensajes [43]. El último documento analizado, utilizó el agrupamiento jerárquico para detectar patrones relacionados a estrategias de aprendizaje y cómo los estudiantes utilizan dichas estrategias [44].

## **Método de Ward (Ward method)**

En agrupación, para el método Ward se han generado varios aportes. Un estudio utilizó el método de Ward para descubrir el rendimiento académico de los estudiantes y analizar el contenido de los cursos a partir de periodos de tiempo [27]. Otro utilizó el mismo método para analiza las estrategias de aprendizaje utilizadas por los estudiantes durante su proceso de formación, agrupando secuencias similares de aprendizaje y

la detección de patrones en el comportamiento [38]. Asimismo, se utilizó el método Ward, con el fin de crear la secuencia de estados de los estudiantes, según el orden cronológico de las sesiones de estudio desarrolladas, analizando el comportamiento derivado de las estrategias y tácticas de estudio [42]. Finalmente se analizó un trabajo que también utiliza la misma técnica para detectar estrategias de gestión del tiempo en las actividades de aprendizaje de los estudiantes [40]. Con respecto al método de vecinos más cercanos, es importante aclarar que los aportes generados por los artículos del corpus de referencias, se encuentran descritos en el respectivo párrafo de la categoría de clasificación.

## **Minería de Datos**

La minería de datos permite comprender relaciones, estructuras y patrones en conjuntos de datos muy grandes. En el análisis solo se encontraron técnicas para minería de datos.

## **Técnicas**

Las técnicas más utilizadas para minería de datos que se identificaron en el análisis de los artículos del corpus de referencias son: análisis de componentes principales, descomposición en valores singulares y minería de datos educativa. A continuación, se describen los principales aportes que se han generado para estas técnicas.

## **Análisis de componentes principales (PCA)**

En un artículo se utilizó PCA, para extraer las conductas de aprendizaje utilizadas por los estudiantes durante el proceso de formación, también para analizar los resultados de aprendizaje a través de los promedios de las calificaciones [32]. Otro estudio propone ACware Tutor, un sistema de tutoría inteligente que incluye PCA para verificar variables de seguimiento relacionadas al proceso de aprendizaje en línea [17]. De la misma forma, otro documento examina los comportamientos de aprendizaje y las características de los estudiantes en una clase de programación para móviles. Crearon visualizaciones de intensidad del trabajo para explorar patrones temporales del comportamiento de los estudiantes y luego los correlacionaron con su rendimiento académico [45]. También se utilizó esta técnica para agrupar los estudiantes de acuerdo a sus compromisos de aprendizaje en línea, analizando sus comportamientos semanales a partir de un sistema de tutoría inteligente basado en ontología [46]. Otro trabajo examinó el efecto de la personalidad con el rendimiento académico, aplicaron PCA para predecir el rendimiento a partir de la personalidad y el nivel de colaboración [19]. También existe una investigación que analiza las diferencias entre los logros de aprendizaje y la participación de los estudiantes con diferentes estilos cognitivos [47]. El último documento analizado utiliza PCA para generar un modelo analítico que se compone de: habilidades técnicas de aprendizaje en línea, gestión de tiempo, apoyo y motivación del estudiante [48].

## Descomposición en valores singulares (SVD)

Durante el análisis de esta técnica se encontraron dos aportes significativos. El primero define un modelo denominado AMoC, que utiliza SVD y técnicas de procesamiento del lenguaje natural para modelar cómo los estudiantes estructuran la información leída. La implementación tiene dos modelos semánticos, que identifican palabras y agregan soporte al grado de similitud con otras palabras [49]. El otro aporte emplea técnicas de procesamiento del lenguaje natural y SVD para identificar automáticamente los vínculos entre las conversaciones por chat con múltiples participaciones, facilitando el seguimiento de los temas y los hilos de discusión [22].

## Minería de datos educativa (EDM)

Se utilizó EDM para analizar el proceso de formación y comprender el comportamiento de aprendizaje de los estudiantes con respecto a la programación [50]. Otros estudios utilizaron esta técnica y la información registrada en el LMS con el fin de comprender las estrategias pedagógicas utilizadas por los estudiantes [51], o en la identificación de estudiantes falsos (estudiantes que utilizan métodos no autorizados para mejorar sus calificaciones) [52], o en el análisis de la influencia que tienen los elementos visuales frente a los compromisos de los estudiantes [53]. Asimismo, para agrupar o explorar los estudiantes de acuerdo a sus comportamientos de aprendizaje en línea y su correlación con el rendimiento académico, permitiendo generar visualizaciones de la intensidad del trabajo del estudiante en función de su rendimiento [17], [45]. Otros documentos utilizan EDM para rastrear los conocimientos previos y hábitos de aprendizaje de los estudiantes [54]. De igual forma se implementó un estudio de mapeo sistemático que analiza las investigaciones existentes sobre la implementación de estrategias y algoritmos de aprendizaje [55]. Otro estudio utiliza un panel interactivo denominado Analyst, que permite a los docentes analizar los datos relacionados con los estudiantes en riesgo de reprobar [56]. También se implementó Experience API para almacenar experiencias de aprendizaje, e identificaron patrones en los hábitos de estudio de los estudiantes [57]. Asimismo, se implementó EDM para identificar patrones que manifiestan actividades no desarrolladas por los estudiantes [58]. Finalmente se analizó un documento que con EDM analizaron la cantidad de veces que el estudiante lee el material antes de tomar un cuestionario, el tiempo dedicado para resolverlo y el comportamiento durante la preparación [15].

En la Tabla 1, se muestran las categorías, los métodos y las técnicas identificadas a partir del corpus de referencias. Los artículos analizados se organizaron en clasificación, regresión, agrupamiento y minería de datos. Se identificaron los métodos (Naive Bayes, árboles de decisión, SVM, proceso gaussiano, vecinos más cercanos, mínimos cuadrados, agrupamiento jerárquico y método Ward) y las técnicas (validación cruzada, aumento del gradiente extremo, regresión lineal, regresión logística, PCA, SVD y EDM).

Tabla 1. Categorías, métodos y técnicas identificadas en el corpus de referencias

Categoría	Método/Técnica	Artículos
Clasificación	Naive Bayes	[8]–[10]
	Árboles de decisión	[8], [9], [11]–[18]
	Máquinas de vectores de soporte (SVM)	[8], [9], [13], [16], [19]–[21]
	Proceso gaussiano	[8], [22]
	Vecinos más cercanos	[13], [21]
	Validación cruzada	[8], [10], [13]–[15], [19], [22]–[24]
	Aumento del gradiente extremo	[8]
Regresión	Mínimos Cuadrados	[25]–[30]
	Árboles de decisión	[8], [9], [11]–[18]
	Máquinas de vectores de soporte (SVM)	[8], [9], [13], [16], [19]–[21]
	Proceso gaussiano	[8], [22]
	Vecinos más cercanos	[13], [21]
	Regresión lineal	[8], [24], [29], [31]–[36]
	Regresión logística	[8], [10], [16], [19], [21], [24], [29], [37]
Agrupación	Agrupamiento jerárquico	[24], [38]–[44]
	Vecinos más cercanos	[13], [21]
	Método de Ward	[27], [38], [40], [42]
Minería de Datos	Análisis de componentes principales (PCA)	[17], [19], [32], [45]–[48]
	Descomposición en valores singulares (SVD)	[22], [49]
	Minería de datos educativa (EDM)	[15], [17], [45], [50]–[58]

Fuente: Elaboración propia

## Metodología

En esta sección, se describe la metodología PRISMA [59], utilizada para realizar la revisión sistemática sobre analítica de aprendizaje, como estrategia de apoyo al aula invertida en los cursos de programación. Incluye las preguntas de investigación que motivaron al desarrollo del trabajo, las fuentes de información, las cadenas de búsqueda, los criterios de inclusión y exclusión, y finalmente el análisis del corpus de referencias.

## Preguntas de investigación

La revisión comenzó con la definición de las preguntas de investigación para el análisis, esto permitió realizar una búsqueda de artículos científicos relacionados al tema objeto de estudio, con el fin de obtener los argumentos correspondientes a cada uno de los interrogantes planteados: ¿Cómo fue el proceso de selección de referencias para el desarrollo de la revisión sistemática?, se responde en esta sección; ¿Qué métodos y técnicas de analítica de aprendizaje sirven como apoyo al aula invertida?, se argumenta en la sección del estado del arte; ¿Cuáles son los aportes de los métodos y técnicas de analítica de aprendizaje,

al aula invertida para los cursos de programación? y ¿Qué herramientas son utilizadas en la analítica de aprendizaje, como apoyo al aula invertida para los cursos de programación?, se responden en la sección de resultados y discusión.

## Fuentes de información

Se realizó la búsqueda de artículos y documentos publicados en revistas y conferencias científicas, con la ayuda de herramientas digitales que permitieron identificar la información para el análisis. Estas fueron: Scopus, ScienceDirect, Web of Science y Google Scholar, cada una de ellas incluye una base de datos que permite documentar la bibliografía correspondiente de los textos que hacen parte del corpus de referencias.

## Cadenas de búsqueda

A continuación, se enumeran las ecuaciones de búsqueda utilizadas para construir el corpus de referencias:

- ("flipped classroom" AND "learning analytics")
- "flipped classroom" AND "learning analytics" (Filter by year: Since 2015)
- ("flipped classroom" AND "artificial intelligence" AND "learning analytics")
- "learning analytics" and "flipped classroom" and "programming course"

En total se hallaron 232 documentos que se encuentran distribuidos de la siguiente manera: Scopus 74 documentos, ScienceDirect 28 documentos, Web of Science 2 documentos y Google Scholar 128 documentos. En las ecuaciones de búsqueda se utilizaron filtros como rango de fechas, subárea de conocimiento y palabras claves, que permitieron mejorar la búsqueda de las referencias.

## Criterios de inclusión y exclusión

Después de identificar todos los artículos de las fuentes de datos, se procedió a eliminar los elementos duplicados con la herramienta Zotero [60, 61, 62], descartando en total de seis documentos. Luego se procedió a realizar la lectura de los resúmenes y filtrar los archivos a partir de los criterios de inclusión definidos:

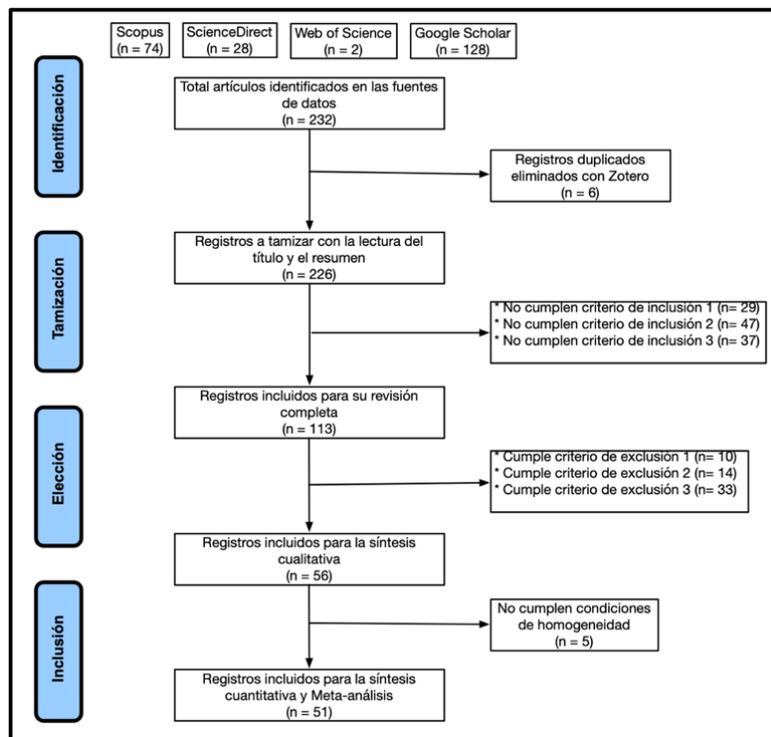
- Artículos que incluyen analítica de aprendizaje y aula invertida.
- Artículos que incluyen métodos y técnicas de analítica de aprendizaje como apoyo al aula invertida.
- Artículos que se relacionan con analítica de aprendizaje, aula invertida y cursos de programación.

Para el primer criterio de inclusión se descartaron 29 artículos del corpus de referencia; en el segundo 47 y en el último 37 documentos. Posteriormente se aplicaron los criterios de exclusión:

- Artículos cuya fecha de publicación es inferior al rango de años definido para el análisis.
- Artículos que solo incluyen analítica de aprendizaje o aula invertida.
- Artículos que incluyen procesos, modelos o algoritmos de analítica de aprendizaje como apoyo al aula invertida.

De los criterios de exclusión se eliminaron 57 artículos, de los cuales 10 corresponden al primer criterio, 14 al segundo criterio y 33 al último criterio. Luego se procedió a descartar 5 artículos de las referencias, porque no cumplen con las condiciones de homogeneidad para el análisis. Finalmente, el corpus de referencias quedó conformado con 51 artículos que sirvieron de apoyo para la síntesis cualitativa, cuantitativa y el meta-análisis (ver Figura 1).

Figura 1. Proceso PRISMA para conformación del corpus de referencias



Fuente: [59]

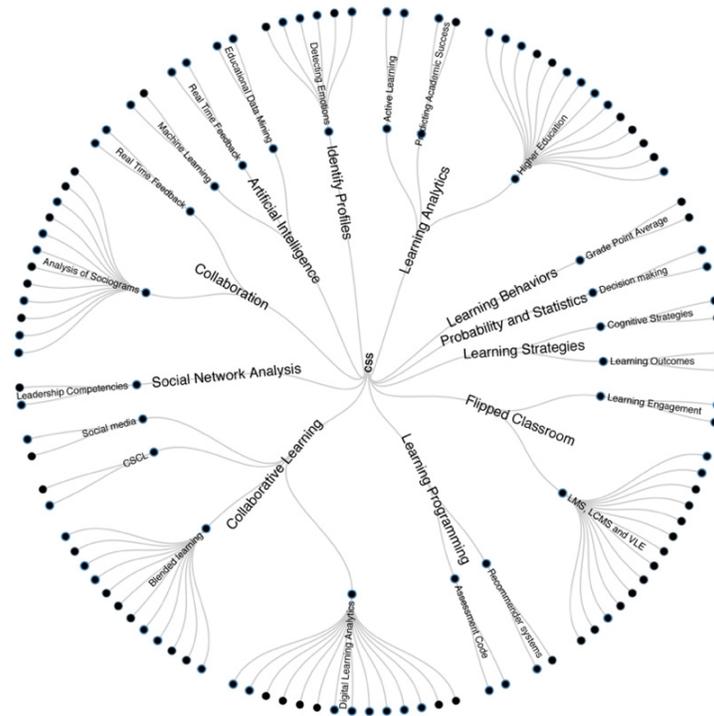
## Análisis del corpus de referencias

Teniendo en cuenta que después de aplicar el proceso Prisma, el corpus de referencias quedó conformado con 51 artículos investigativos, se procedió a realizar el análisis de los documentos encontrados por año de publicación; áreas y campos de interés; y tipos de documentos, con el fin de dar respuesta a la pregunta: ¿Cuáles son las áreas de conocimiento identificadas a partir del corpus de referencias?

Para la revisión sistemática de literatura se definió como rango de búsqueda los años comprendidos entre 2015 y 2020. Con base a esta información, la cantidad de documentos seleccionados por año fue: 3 artículos para 2015; 3 artículos para 2016; 2 artículos para 2017; 28 artículos para 2018; 12 artículos para 2019 y 3 artículos para 2020. La participación en términos porcentuales de los artículos para la revisión sistemática de literatura fue: 3.92% para el año 2017; 5.88% para los años 2015, 2016 y 2020; 23.53% para el año 2019 y 54.91% para el año 2018.

Luego se identificaron las áreas de conocimiento, a partir del Sistema de Clasificación de Computación (CCS) definido por ACM. En la Figura 2 se pueden observar 11 áreas de conocimiento que fueron reconocidas a partir de los documentos seleccionados en el corpus de referencia.

Figura 2. Clasificación del corpus de referencias por área de conocimiento, según clasificación CSS de ACM



Fuente: Elaboración propia

Todas estas áreas de conocimiento han generado aportes significativos en la educación y sus estrategias de aprendizaje. Adicionalmente han incorporado en sus procesos métodos, técnicas, modelos y algoritmos. Debido a la cantidad de información por analizar, se decidió trabajar solo con los métodos y las técnicas que hacen parte del área de la inteligencia artificial y las subáreas: aprendizaje automático y minería de datos. Encontrando que el aprendizaje automático se ha utilizado para solucionar problemas relacionados al aprendizaje supervisado y no supervisado. Con respecto al aprendizaje supervisado se analizaron los métodos y las técnicas correspondientes a los campos de interés de clasificación y regresión; para el aprendizaje no supervisado se analizaron los métodos de agrupación y en minería de datos se analizó EDM y otras técnicas.

La cantidad de artículos analizados por campo de interés, para el desarrollo de este documento fue: 18 artículos para clasificación; 13 artículos para regresión; 6 artículos para agrupación y 14 artículos para minería de datos. En términos porcentuales para el campo de interés de clasificación se analizó el 35.29% del total de los artículos seleccionados; regresión se analizó el 25.49%; agrupación el 11.77% y minería de datos el 27.45%. Asimismo, los tipos de documentos que constituyen el corpus se clasifican en: partes o capítulos de libros (27 documentos), artículos (19 documentos) y conferencias (5 documentos). Esto indica que el 52.94% de los documentos corresponden a partes o capítulos de libros; el 37.26% artículos y el 9.80% a conferencias.

## Resultados y discusión

En esta sección se responden las preguntas: ¿Cuáles son los aportes de los métodos y técnicas de analítica de aprendizaje, al aula invertida para los cursos de programación? y ¿Qué herramientas son utilizadas en la analítica de aprendizaje, como apoyo al aula invertida para los cursos de programación?, finalmente se encuentra la discusión.

### Aportes para los cursos de programación

Con base al corpus de referencias se encontraron varios aportes importantes para los cursos de programación. Las contribuciones identificadas durante el análisis, se organizaron por año, con el fin de visualizar los métodos y las técnicas que se han utilizado en la generación de dichos aportes. También para observar los elementos y objetivos definidos en el desarrollo de las investigaciones.

En los años 2015, 2017 y 2018 se encontraron seis aportes que incluyen los métodos y las técnicas: validación cruzada, árboles de decisión, mínimos cuadrados, PCA y EDM. Se analizaron las características, la participación y el comportamiento de los estudiantes en cursos de programación: Java, desarrollo móvil, estructuras de datos y algoritmos, con el fin de identificar patrones de comportamiento de los estudiantes y luego correlacionarlos con su rendimiento académico [25, 45, 50]. También se analizaron las estrategias de programación utilizadas por los estudiantes para el desarrollo de talleres que incluyen codificación, verificando si existe una relación entre dichas estrategias y la solución de actividades propuestas durante el desarrollo del curso [23]. Del mismo modo se analizó la relación entre la preparación de los estudiantes antes de clase y los puntajes obtenidos en las pruebas, en cursos de programación con C [15]. Asimismo, otro aporte hace referencia a un sistema de aprendizaje automático, desarrollado para detectar los errores más comunes generados por los estudiantes durante el proceso de codificación, el software al final genera los respectivos comentarios y la retroalimentación al aprendiz [14].

Durante los años 2019 y 2020 se encontraron contribuciones que incluyen los métodos y las técnicas: mínimos cuadrados, regresión logística, EDM, árboles de decisión, PCA y regresión lineal. Existe un documento que analiza las herramientas utilizadas para el aprendizaje de la programación a partir de juegos, con el fin de ayudar a mejorar la comprensión de los conceptos de programación y las experiencias de aprendizaje de los estudiantes [28]. También se identificó una investigación que analiza las diferencias entre las actividades en línea y las experiencias de aprendizaje en un curso de programación, el objetivo del estudio se centró en priorizar las necesidades de apoyo de los estudiantes y la identificación temprana de aprendices en riesgo [37]. Del mismo modo se encontró un trabajo que rastrea las características de los estudiantes, los conocimientos previos y hábitos de aprendizaje, para generar información relacionada al proceso de aprendizaje y el abandono de los estudiantes [54].

También se encontró un trabajo que define un método basado en reglas, que permite corregir de manera automática los errores cometidos por los estudiantes durante la composición de un lenguaje [17]. Finalmente se encontró una investigación que evalúa los efectos del aula invertida en la enseñanza de la ingeniería de software, a partir de las calificaciones generadas en LMS, el objetivo de la investigación se basó en analizar los efectos de cambiar la evaluación tradicional al aula invertida, descubriendo que al aplicar el FC, se aumentan las calificaciones y los estudiantes mejoran las habilidades de aprendizaje [31].

En la Tabla 2, se encuentran los aportes identificados durante el análisis de la información, que contribuyen a los cursos de programación. Cada uno de los elementos se organizó por año, métodos y técnicas que incluye el aporte, y finalmente su descripción. Presentar la información con este orden, permite observar las características y objetivos definidos para el desarrollo de las investigaciones.

**Tabla 2. Aportes de los métodos y técnicas de analítica de aprendizaje, al aula invertida para los cursos de programación.**

Año	Método/Técnica	Aporte
2015	Validación cruzada, árboles de decisión, EDM	Análisis de la relación entre la preparación de los estudiantes antes de clase y los puntajes obtenidos en las pruebas, en cursos de programación en C [15].
2017	Mínimos cuadrados	Análisis de la participación y el comportamiento de los estudiantes, en cursos de programación Java (básicos y avanzados) con actividades de resolución de problemas [25].
2018	Validación cruzada	Estrategias de programación utilizadas por el estudiante para desarrollar talleres de codificación y verificar la relación existente entre las habilidades implementadas y la solución de las actividades propuestas [23].
	Validación cruzada, árboles de decisión	Sistema de aprendizaje automático que permite detectar los errores más comunes de codificación y proporciona los respectivos comentarios para el estudiante [14].
	PCA, EDM	Análisis de los comportamientos de aprendizaje y las características de los estudiantes en un curso de programación para móviles [45].
	EDM	Registrar los comportamientos de los estudiantes y analizar el impacto de aprendizaje asociado al monitoreo y progreso de cursos de estructuras de datos y algoritmos [50].
2019	Mínimos cuadrados	Análisis de las herramientas utilizadas para el aprendizaje de la programación a partir de juegos [28].
	Regresión logística	Análisis de diferencias entre las actividades en línea y la experiencia de aprendizaje, utilizando datos de un curso de programación [37].
	EDM	Rastreo de las características de los estudiantes, relacionadas a conocimientos previos y hábitos de aprendizaje, en cursos de ciencias de la computación [54].
2020	Árboles de decisión, PCA, EDM	Método basado en reglas, que permite corregir de manera automática los errores cometidos por los estudiantes en la composición de un lenguaje [17].
	Regresión lineal	Evaluación de los efectos del aula invertida en la enseñanza de la ingeniería de software a partir de las calificaciones [31].

Fuente: Elaboración propia

## Herramientas para los cursos de programación

Con el análisis de los documentos seleccionados en el corpus de referencias, se encontraron varias herramientas que se utilizaron durante las investigaciones y contribuyen a los cursos de programación. Cada uno de los aportes se organizó por los métodos y las técnicas donde se incluye la herramienta, el nombre de la herramienta y finalmente su descripción (ver Tabla 3). Las contribuciones más importantes de las herramientas identificadas son: reconocimiento de las emociones textuales de los estudiantes,

comportamiento de los estudiantes, experiencias de aprendizaje, anotación y consumo de videoconferencias, herramientas de programación basadas en juegos y herramientas para la comprensión de conceptos.

**Tabla 3. Herramientas utilizadas en la analítica de aprendizaje, como apoyo al aula invertida para los cursos de programación.**

	Método/Técnica	Herramienta	Descripción
1	Naive Bayes, árboles de decisión, SVM	EmoTect	Es un sistema que reconoce las emociones textuales de los estudiantes, sirve para la toma de decisiones en la prestación de asesorías [9].
2	Árboles de decisión, EDM	Experience API	Es un software que permite almacenar las interacciones y experiencias de aprendizaje de los estudiantes, de las diferentes actividades de formación en línea [57].
3	Árboles de decisión, PCA, EDM	AC-ware Tutor	Es un sistema de tutoría inteligente, también se utiliza para analizar el comportamiento de los estudiantes en entornos LMS [17, 46].
4	Árboles de decisión, SVM, vecinos más cercanos, validación cruzada	LIWC	Es un software que analiza textos a través de consultas lingüísticas y recuento de palabras [13].
5	SVM, vecinos más cercanos, regresión logística	AttentiveLearner	Es un sistema inteligente de aprendizaje móvil, optimizado para consumir videos de conferencias en MOOC [21].
6	Regresión lineal	CLAS	Sistema de anotación de lectura colaborativa, una aplicación de anotación de video basada en la web [36].
7	Mínimos cuadrados	GBPT	Herramientas de programación basadas en juegos, ayuda a comprender los conceptos de programación a través de una experiencia de aprendizaje agradables [28].
8	Agrupamiento jerárquico, descomposición en valores singulares	ReaderBench	Es un marco de trabajo que ayuda a la predicción y evaluación de la comprensión (materiales, comportamiento, explicaciones metacognitivas) a través de técnicas de minería de texto, procesamiento del lenguaje natural y herramientas de análisis de redes sociales [43, 49].
9	Descomposición en valores singulares	AMoC	Modelo automatizado de comprensión, utiliza técnicas avanzadas de procesamiento del lenguaje natural, para modelar como los estudiantes conceptualizan la información leída [49].
10	EDM	WebPGA	Es una plataforma que conecta los espacios de aprendizaje físico y digital en el aprendizaje de programación [50].
11	EDM	Analyst	Panel interactivo basado en la web, brinda al docente la oportunidad de analizar datos relacionados con los estudiantes de diversas fuentes [56].

**Fuente: Elaboración propia**

Luego de generar la revisión sistemática y analizar los aportes generados por LA como apoyo al FC, estableciendo como objetivo los métodos, técnicas y herramientas que han apoyado al aprendizaje de la programación; es importante aclarar que a la fecha no se ha desarrollado un modelo de predicción basado en analítica de aprendizaje y aula invertida, que permita identificar las competencias que requiere un programador para desempeñar cargos en la industria del software. De esta manera se puede cerrar la brecha existente entre las necesidades del sector productivo y el sector educativo que están relacionadas con la programación.

Como trabajo futuro se propone trabajar en las evaluaciones presentadas por los estudiantes antes de la clase, para cursos de programación con aula invertida. Con el fin de identificar los conocimientos adquiridos por el aprendiz, durante la revisión de los materiales y videos extra clases que se deben desarrollar en sistemas de gestión de aprendizaje (LMS). De esta forma se puede contribuir a las inquietudes presentadas por los estudiantes, ayudando al cumplimiento de los objetivos propuestos para el curso y disminuyendo la tasa de abandono.

## Conclusiones

Este trabajo presenta una revisión sistemática de literatura, de los métodos y las técnicas de LA que sirven de apoyo al FC. Cada una de las contribuciones han aportado a las áreas de conocimiento de aprendizaje automático y minería de datos, con la solución de problemas relacionados al aprendizaje supervisado, no supervisado y minería de datos educativa. Las investigaciones analizadas se han enfocado en el estudio de los datos almacenados en los sistemas de gestión de aprendizaje (LMS), para generar modelos de predicción que permitan mejorar los comportamientos y hábitos de estudio, el rendimiento académico, las tutorías y la retención de los estudiantes.

Los aportes más significativos encontrados en los métodos y las técnicas de los cursos de programación, se han orientado en la preparación de los estudiantes para las evaluaciones, la definición de estrategias en el aprendizaje de la programación, el desarrollo de sistemas que permiten la detección de errores en el código fuente y sistemas de tutorías inteligentes. Todas estas contribuciones han ayudado a mejorar la retención de los estudiantes, los comportamientos de aprendizaje en línea, la solución de problemas con trabajo colaborativo y la retroalimentación personalizada.

Las herramientas identificadas como apoyo a los cursos de programación han contribuido con software que permite reconocer las emociones textuales de los estudiantes, sistemas de tutoría inteligentes y aplicaciones para consumir videos optimizados en ambientes con aula invertida. Esto ha permitido mejorar en los estudiantes su participación, las experiencias de aprendizaje y el comportamiento en entornos LMS; también han apoyado a los docentes en la toma de decisiones frente al desarrollo de los cursos y la retroalimentación personalizada de los aprendices.

## Referencias bibliográficas

1. "El e-Learning dominará la educación a partir de 2019 | VIU." <https://www.universidadviu.com/el-e-learning-dominara-la-educacion-a-partir-de-2019/> (accessed Jul. 14, 2020).
2. G. Siemens, "Learning Analytics: The Emergence of a Discipline," *Am. Behav. Sci.*, vol. 57, no. 10, pp. 1380–1400, Oct. 2013, DOI: 10.1177/0002764213498851.
3. P. Rojas Castro, "Learning Analytics. Una Revisión de la Literatura," *Educ. Educ.*, vol. 20, no. 1, pp. 106–128, Feb. 2017, DOI: 10.5294/edu.2017.20.1.6.
4. J. P. Campbell, "Academic Analytics: A New Tool for a New Era," p. 10, 2007.
5. S. Nunn, J. T. Avella, T. Kanai, and M. Kebritchi, "Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review," *Online Learn.*, vol. 20, no. 2, Jan. 2016, DOI:10.24059/olj.v20i2.790.

6. J. L. Bishop, U. S. University, D. M. A. Verleger, E.-R. A. Univ, and D. Beach, "The Flipped Classroom: A Survey of the Research," vol. 30, pp. 1–18, 2013.
7. H. N. Mok, "Teaching tip: The flipped classroom," *J. Inf. Syst. Educ.*, vol. 25, pp. 7–11, 2014.
8. A. Y. Q. Huang, O. H. T. Lu, J. C. H. Huang, C. J. Yin, and S. J. H. Yang, "Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs," *Interact. Learn. Environ.*, vol. 28, no. 2, pp. 206–230, Feb. 2020, DOI: 10.1080/10494820.2019.1636086.
9. E. A. Kolog, "Detecting Emotions in Students' Generated Content: An Evaluation of EmoTect System," in *Technology in Education. Innovative Solutions and Practices*, Singapore, 2018, vol. 843, pp. 235–248, DOI: [https://doi.org/10.1007/978-981-13-0008-0\\_22](https://doi.org/10.1007/978-981-13-0008-0_22).
10. F. Ghaffar, N. Peirce, and A. Serlie, "Assessing Leadership Competencies Through Social Network Analysis," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 452–466, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_35](https://doi.org/10.1007/978-3-319-98572-5_35).
11. L. Wang and T. Wang, "Automatic Correction of Definite Article Redundancy Error in the English Compositions of College Students," in *Technology in Education. Innovative Solutions and Practices*, Singapore, 2018, vol. 843, pp. 77–85, DOI: [https://doi.org/10.1007/978-981-13-0008-0\\_8](https://doi.org/10.1007/978-981-13-0008-0_8).
12. Mbaye Baba, "Organisation of knowledge from traces of human learning," Portugal, 2019, DOI: 10.33965/el2019\_201909D055.
13. V. Neto *et al.*, "Automated Analysis of Cognitive Presence in Online Discussions Written in Portuguese," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 245–261, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_19](https://doi.org/10.1007/978-3-319-98572-5_19).
14. B. Burke, P. Weßeler, and J. te Vrugt, "A Programming Language Independent Platform for Algorithm Learning," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 652–655, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_66](https://doi.org/10.1007/978-3-319-98572-5_66).
15. Y. Yamada and M. Hirakawa, "A Case Study of Analyzing Logs of LMS in Flipped Classroom," in *2015 IIAI 4th International Congress on Advanced Applied Informatics*, Okayama, Japan, Jul. 2015, pp. 374–378, DOI: 10.1109/IIAI-AAI.2015.255.
16. M. C. Saiz Manzanares, R. Marticorena Sánchez, Á. Arnaiz González, M. del C. Escolar Llamazares, and M. Á. Queiruga Dios, "Detección del alumno en riesgo en titulaciones de Ciencias de la Salud: aplicación de técnicas de Learning Analytics," *Eur. J. Investig. Health Psychol. Educ.*, vol. 8, no. 3, p. 129, Dec. 2018, DOI: 10.30552/ejihpe.v8i3.273.
17. I. Šarić-Grgić, A. Grubišić, L. Šerić, and T. J. Robinson, "Student Clustering Based on Learning Behavior Data in the Intelligent Tutoring System:," *Int. J. Distance Educ. Technol.*, vol. 18, no. 2, pp. 73–89, Apr. 2020, DOI: 10.4018/IJDET.2020040105.
18. S. Nizam Ismail, S. Hamid, and H. Chiroma, "The utilization of learning analytics to develop student engagement model in learning management system," *J. Phys. Conf. Ser.*, vol. 1339, p. 012096, Dec. 2019, DOI: 10.1088/1742-6596/1339/1/012096.
19. M. Rahmani Hanzaki and C. Demmans Epp, "The Effect of Personality and Course Attributes on Academic Performance in MOOCs," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 497–509, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_38](https://doi.org/10.1007/978-3-319-98572-5_38).
20. S. Praharaj, M. Scheffel, H. Drachsler, and M. Specht, "Multimodal Analytics for Real-Time Feedback in Co-located Collaboration," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 187–201, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_15](https://doi.org/10.1007/978-3-319-98572-5_15).

21. P. Pham and J. Wang, "Attentivelearner: Improving mobile MOOC learning via implicit heart rate tracking," *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.*, vol. 9112, pp. 367–376, 2015, DOI: 10.1007/978-3-319-19773-9\_37.
22. M. Masala, S. Ruseti, G. Gutu-Robu, T. Rebedea, M. Dascalu, and S. Trausan-Matu, "Help Me Understand This Conversation: Methods of Identifying Implicit Links Between CSCL Contributions," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 482–496, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_37](https://doi.org/10.1007/978-3-319-98572-5_37).
23. K. Sharma, K. Mangaroska, H. Trætterberg, S. Lee-Cultura, and M. Giannakos, "Evidence for Programming Strategies in University Coding Exercises," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 326–339, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_25](https://doi.org/10.1007/978-3-319-98572-5_25).
24. J. Maldonado-Mahauad, M. Pérez-Sanagustín, P. M. Moreno-Marcos, C. Alario-Hoyos, P. J. Muñoz-Merino, and C. Delgado-Kloos, "Predicting Learners' Success in a Self-paced MOOC Through Sequence Patterns of Self-regulated Learning," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 355–369, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_27](https://doi.org/10.1007/978-3-319-98572-5_27).
25. F. H. Wang, "An exploration of online behaviour engagement and achievement in flipped classroom supported by learning management system," *Comput. Educ.*, vol. 114, pp. 79–91, Nov. 2017, DOI: 10.1016/j.compedu.2017.06.012.
26. Z. Tao, B. Zhang, and I. K. W. Lai, "Perceived Online Learning Environment and Students' Learning Performance in Higher Education: Mediating Role of Student Engagement," in *Technology in Education. Innovative Solutions and Practices*, Singapore, 2018, vol. 843, pp. 56–64, DOI: [https://doi.org/10.1007/978-981-13-0008-0\\_6](https://doi.org/10.1007/978-981-13-0008-0_6).
27. Z. Sun, L. Lu, and K. Xie, "The Effects of Self-Regulated Learning on Students' Performance Trajectory in the Flipped Math Classroom," *Singap. Int. Soc. Learn. Sci.*, p. 8, 2016, DOI: 10.22318/icsl2016.11.
28. Su Chien-Yuan, Hu Yue, and Li Yu-Hang, "Teacher attitudes regarding the use of game-based programming tools in K-12 education," in *MULTI CONFERENCE ON COMPUTER SCIENCE AND INFORMATION SYSTEMS 2019*, Portugal, 2019, p. 73, DOI: 10.33965/el2019\_201909F010.
29. D. Gašević, S. Dawson, T. Rogers, and D. Gasevic, "Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success," *Internet High. Educ.*, vol. 28, pp. 68–84, Jan. 2016, DOI: 10.1016/j.iheduc.2015.10.002.
30. C.-H. Luk, K.-K. Ng, and W.-M. Lam, "The Acceptance of Using Open-Source Learning Platform (Moodle) for Learning in Hong Kong's Higher Education," in *Technology in Education. Innovative Solutions and Practices*, Singapore, 2018, vol. 843, pp. 249–257, DOI: [https://doi.org/10.1007/978-981-13-0008-0\\_23](https://doi.org/10.1007/978-981-13-0008-0_23).
31. L. Gren, "A Flipped Classroom Approach to Teaching Empirical Software Engineering," *IEEE Trans. Educ.*, 2020, DOI: 10.1109/TE.2019.2960264.
32. C. C. Hsiao, J. C. H. Huang, A. Y. Q. Huang, O. H. T. Lu, C. J. Yin, and S. J. H. Yang, "Exploring the effects of online learning behaviors on short-term and long-term learning outcomes in flipped classrooms," *Interact. Learn. Environ.*, vol. 27, no. 8, pp. 1160–1177, 2019, DOI: 10.1080/10494820.2018.1522651.
33. I. Khan and A. Pardo, "Data2U: scalable real time student feedback in active learning environments," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, Edinburgh, United Kingdom, 2016, pp. 249–253, DOI:10.1145/2883851.2883911.
34. H. Le and J. Jia, "Analysis of Learner Timeout Behavior in Online Tests of a Bigdata Set Based on the OLAI Concept," in *Technology in Education. Innovative Solutions and Practices*, Singapore, 2018, vol. 843, pp. 285–294, DOI: [https://doi.org/10.1007/978-981-13-0008-0\\_27](https://doi.org/10.1007/978-981-13-0008-0_27).

35. J. Jia, "Design, Implementation and Evaluation of Blended Learning for the Undergraduate Course 'Education and Artificial Intelligence,'" in *Technology in Education. Innovative Solutions and Practices*, Singapore, 2018, vol. 843, pp. 211–222, DOI: [https://doi.org/10.1007/978-981-13-0008-0\\_20](https://doi.org/10.1007/978-981-13-0008-0_20).
36. A. Pardo, Y. Zhao, N. Mirriahi, A. Zhao, S. Dawson, and D. Gašević, "Identifying learning strategies associated with active use of video annotation software," in *ACM International Conference Proceeding Series*, 2015, vol. 16-20-March-2015, pp. 255–259, DOI: 10.1145/2723576.2723611.
37. P. Schwarzenberg, J. Navon, and M. Pérez-Sanagustín, "Models to provide guidance in flipped classes using online activity," *J. Comput. High. Educ.*, 2019, DOI: 10.1007/s12528-019-09233-y.
38. J. Jovanović, D. Gašević, S. Dawson, A. Pardo, and N. Mirriahi, "Learning analytics to unveil learning strategies in a flipped classroom," *Internet High. Educ.*, vol. 33, pp. 74–85, 2017, DOI: 10.1016/j.iheduc.2017.02.001.
39. J. Saint, D. Gašević, and A. Pardo, "Detecting Learning Strategies Through Process Mining," *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.*, vol. 11082 LNCS, pp. 385–398, 2018, DOI: 10.1007/978-3-319-98572-5\_29.
40. N. Ahmad Uzir, D. Gašević, W. Matcha, J. Jovanović, and A. Pardo, "Analytics of time management strategies in a flipped classroom," *J. Comput. Assist. Learn.*, vol. 36, no. 1, pp. 70–88, 2020, DOI:10.1111/jcal.12392.
41. Z. Sun, L. Lu, and K. Xie, "The effects of self-regulated learning on students' performance trajectory in the flipped math classroom," *Proc. Int. Conf. Learn. Sci. ICLS*, vol. 1, pp. 66–73, 2016, DOI: 10.22318/icls2016.11.
42. E. Fincham, D. Gašević, J. Jovanović, and A. Pardo, "From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations," *IEEE Trans. Learn. Technol.*, vol. 12, no. 1, pp. 59–72, 2019, DOI: 10.1109/TLT.2018.2823317.
43. M. Dascalu, M.-D. Sirbu, G. Gutu-Robu, S. Ruseti, S. A. Crossley, and S. Trausan-Matu, "Cohesion-Centered Analysis of Sociograms for Online Communities and Courses Using ReaderBench," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 622–626, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_59](https://doi.org/10.1007/978-3-319-98572-5_59).
44. W. Matcha, D. Gašević, N. A. Uzir, J. Jovanović, and A. Pardo, "Analytics of learning strategies: Associations with academic performance and feedback," in *ACM International Conference Proceeding Series*, 2019, pp. 461–470, DOI: 10.1145/3303772.3303787.
45. A. AlJarrah, M. K. Thomas, and M. Shehab, "Investigating temporal access in a flipped classroom: procrastination persists," *Int. J. Educ. Technol. High. Educ.*, vol. 15, no. 1, 2018, DOI: 10.1186/s41239-017-0083-9.
46. I. Šarić, A. Grubišić, L. Šerić, and T. J. Robinson, "Data-Driven Student Clusters Based on Online Learning Behavior in a Flipped Classroom with an Intelligent Tutoring System," in *Intelligent Tutoring Systems*, Cham, 2019, vol. 11528, pp. 72–81, DOI: [https://doi.org/10.1007/978-3-030-22244-4\\_10](https://doi.org/10.1007/978-3-030-22244-4_10).
47. F.-R. Sun, H.-Z. Hu, R.-G. Wan, X. Fu, and S.-J. Wu, "A learning analytics approach to investigating pre-service teachers' change of concept of engagement in the flipped classroom," *Interact. Learn. Environ.*, 2019, DOI: 10.1080/10494820.2019.1660996.
48. M. Henderikx, K. Kreijns, and M. Kalz, "A Classification of Barriers that Influence Intention Achievement in MOOCs," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 3–15, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_1](https://doi.org/10.1007/978-3-319-98572-5_1).

49. M. Dascalu, I. C. Paraschiv, D. S. McNamara, and S. Trausan-Matu, "Towards an Automated Model of Comprehension (AMoC)," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 427–436, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_33](https://doi.org/10.1007/978-3-319-98572-5_33).
50. Y. V. Paredes, D. Azcona, I.-H. Hsiao, and A. Smeaton, "Learning by Reviewing Paper-Based Programming Assessments," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 510–523, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_39](https://doi.org/10.1007/978-3-319-98572-5_39).
51. J. C. S. Silva, E. Zambom, R. L. Rodrigues, J. L. C. Ramos, and F. Da Fonseca De Souza, "Effects of learning analytics on students' self-regulated learning in flipped classroom," *Int. J. Inf. Commun. Technol. Educ.*, vol. 14, no. 3, pp. 91–107, 2018, DOI: 10.4018/IJICTE.2018070108.
52. G. Alexandron, J. A. Ruipérez-Valiente, S. Lee, and D. E. Pritchard, "Evaluating the Robustness of Learning Analytics Results Against Fake Learners," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 74–87, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_6](https://doi.org/10.1007/978-3-319-98572-5_6).
53. B. Yousuf, A. Staikopoulos, and O. Conlan, "Motivating Students to Enhance Their Knowledge Levels Through Personalized and Scrutable Visual Narratives," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 136–150, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_11](https://doi.org/10.1007/978-3-319-98572-5_11).
54. M. Apiola, T. Karunaratne, E. Kaila, and M.-J. Laakso, "Experiences from Digital Learning Analytics in Finland and Sweden: A Collaborative Approach," in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, May 2019, pp. 627–632, DOI: 10.23919/MIPRO.2019.8757204.
55. E. Doko and L. A. Bexheti, "A systematic mapping study of educational technologies based on educational data mining and learning analytics," in *2018 7th Mediterranean Conference on Embedded Computing, MECO 2018 - Including ECYPS 2018, Proceedings*, 2018, pp. 1–4, DOI: 10.1109/MECO.2018.8406052.
56. J. Vaclavek, J. Kuzilek, J. Skocilas, Z. Zdrahal, and V. Fuglik, "Learning Analytics Dashboard Analysing First-Year Engineering Students," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 575–578, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_48](https://doi.org/10.1007/978-3-319-98572-5_48).
57. A. Gelan *et al.*, "Affordances and limitations of learning analytics for computer-assisted language learning: a case study of the VITAL project," *Comput. Assist. Lang. Learn.*, vol. 31, no. 3, pp. 294–319, 2018, DOI: 10.1080/09588221.2017.1418382.
58. S. Sosnovsky, L. Müter, M. Valkenier, M. Brinkhuis, and A. Hofman, "Detection of Student Modelling Anomalies," in *Lifelong Technology-Enhanced Learning*, Cham, 2018, vol. 11082, pp. 531–536, DOI: [https://doi.org/10.1007/978-3-319-98572-5\\_41](https://doi.org/10.1007/978-3-319-98572-5_41).
59. D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Int. J. Surg.*, vol. 8, no. 5, pp. 336–341, Jan. 2010, DOI: 10.1016/j.ijsu.2010.02.007.
60. R. Avellaneda, S. Cabrera, P. A. Martínez, y C. G. Donoso Albarracín, "Apoyo tecnológico para la fidelización y captación de nuevos clientes por medio de una aplicación móvil," *Investigación e Innovación en Ingenierías*, vol. 5, n.º 1, pp. 92–101, 2017. DOI: <https://doi.org/10.17081/invinno.5.1.2618>
61. J. A. Calderón Velasco, G. A. Amarillo Cárdenas, L. A. Silva Bahamon, y C. G. Donoso Albarracín, "Biometría dactilar: una nueva alternativa de controlar efectivamente la asistencia a clases," *Investigación e Innovación en Ingenierías*, vol. 6, n.º 1, pp. 27 - 39, 2017. DOI: <https://doi.org/10.17081/invinno.6.1.2773>
62. ["Zotero | Your personal research assistant." <https://www.zotero.org/> (accessed Jun. 23, 2020).