

Minería de Datos Educativos: una visión holística

Educational Data Mining: an holistic view

Oswaldo Moscoso-Zea
Facultad de Ciencias de Ingeniería
Universidad Tecnológica Equinoccial
Quito, Ecuador
omoscso@ute.edu.ec

Sergio Luján-Mora
Departamento de Lenguajes y Sistemas Informáticos
Universidad de Alicante
Alicante, España
sergio.lujan@ua.es

Resumen — La minería de datos (MD) agrupa un amplio número de técnicas y algoritmos que permiten extraer conocimiento de bases de datos para la toma de decisiones oportunas. La MD se ha aplicado a diferentes campos siendo uno de ellos el educativo. La aplicación de MD en el campo educativo se conoce como minería de datos educacionales (MDE). El propósito fundamental de MDE es analizar datos de instituciones educativas mediante el uso de diferentes técnicas como: predicción, *clustering*, análisis de series de tiempo, clasificación, entre otras. Este artículo presenta una visión holística de MDE que abarca una clasificación de los algoritmos, métodos y herramientas usados en procesos de MD y un análisis de procesos e indicadores con potencial de mejoras en instituciones educativas. Este estudio abarca artículos presentados desde el año 2005 hasta el año 2015.

Minería de Datos Educativos; Gestión Educativa, Analítica de Datos; Gestión del Conocimiento.

Abstract — Datamining (DM) brings together a wide range of techniques and algorithms which allow the extraction of knowledge from databases for timely decision making. DM has been applied to different fields of study. One important research field is Education. Applying DM in education is known as educational datamining (EDM). The main purpose of EDM is to analyze data from educational institutions using different techniques such as: prediction, clustering, time-series analysis, classification, among others. This paper presents an holistic view of EDM including classification of algorithms, methods and tools used in DM processes. Furthermore, processes and indicators that could be improved are analyzed in educational institutions. This study covers papers presented from 2005 to 2015.

Educational DataMining; Educational Management; Data Analytics; Knowledge Management.

I. INTRODUCCIÓN

La minería de datos (MD) pretende resolver problemas complejos cuya solución no se puede hallar con técnicas tradicionales como la estadística. Esto se logra al descubrir patrones y predecir tendencias por medio del análisis de datos generados a través de los diferentes sistemas operacionales y transaccionales de una institución y que están almacenados en sus bases de datos [1]. Cuando se aplica minería de datos en instituciones educativas la disciplina se conoce como minería de datos educacionales (MDE).

La MDE es una disciplina en evolución que usa tecnologías informáticas como son almacenes de datos y herramientas de inteligencia de negocios para descubrir tendencias y patrones

sobre datos educacionales. El conocimiento que la MD genera apoya a las autoridades de centros de educación superior en la toma de decisiones oportunas y a los profesores para analizar el comportamiento y aprendizaje de sus alumnos [2]. La disciplina se enfoca en el diseño de modelos para mejorar las experiencias del aprendizaje y la eficiencia organizacional [3].

La MD es una tarea importante del proceso de descubrimiento de conocimiento en bases de datos conocido en inglés como *Knowledge Discovery in Databases* (KDD). KDD busca descubrir patrones y tendencias con la información almacenada en sus repositorios de datos [4]. Un proceso KDD puede ayudar a que una organización o institución educativa gestione sus recursos de una manera eficiente y pueda adquirir ventaja competitiva [5].

En este artículo se presenta una revisión bibliográfica que incluye artículos sobre MD y MDE desde el año 2005 hasta el año 2015. Los autores de los artículos revisados han analizado la disciplina de una forma puntual y específica, por ejemplo se analiza de forma separada procesos educativos, plataformas educativas virtuales, servicios y desempeño organizacional de instituciones educativas, sin analizar la disciplina de forma global. Además, durante la presente investigación se ha detectado que existe mucha confusión en los algoritmos y métodos de MD aplicables a MDE. Por otro lado, existe poca bibliografía de los indicadores de gestión que se pueden mejorar en las instituciones educativas.

Este artículo propone analizar las distintas posibilidades que ofrece la MDE y presentar una visión holística de la disciplina. La palabra holística plantea la concepción de una realidad como un todo, distinto de la suma de las partes que lo componen. En este artículo la visión holística se interpreta como un análisis global de la disciplina. Este análisis permite resolver los problemas detectados, proponiendo un orden para clasificar los algoritmos, métodos y herramientas. Además, se detallan los posibles escenarios para experimentar con MDE.

La estructura de este artículo es la siguiente: después de introducir el tema en esta sección, continuaremos en la sección 2 explicando los fundamentos de MD y de MDE y ejemplificando instituciones que se han beneficiado del uso de MDE. En la sección 3 presentaremos los resultados de la investigación incluyendo los algoritmos y métodos más usados por los investigadores. Además, se presentará un análisis de los contextos donde se ha experimentado con MDE para detectar

escenarios potenciales de experimentación. Finalmente, las conclusiones de la investigación se presentarán en la sección 4.

II. MINERÍA DE DATOS

El término “minería de datos” fue acuñado a inicios de los años 90 por la comunidad de investigadores de base de datos. También es conocido como “Arqueología de Datos”, “Recolección de Datos”, “Extracción de Conocimiento”, “Analítica de Datos” y es parte fundamental del proceso de descubrimiento de conocimiento en bases de datos KDD [6]. La MD es el proceso de emplear tecnologías de información para analizar y extraer conocimiento de la información contenida en repositorios de datos de la organización. El marco de referencia más usado para entender el ciclo de vida de un proyecto de MD es CRISP-DM, las fases principales como se muestra en la Figura 1 son: comprensión del negocio, comprensión de los datos, preparación de los datos, creación de modelos, evaluación y despliegue.

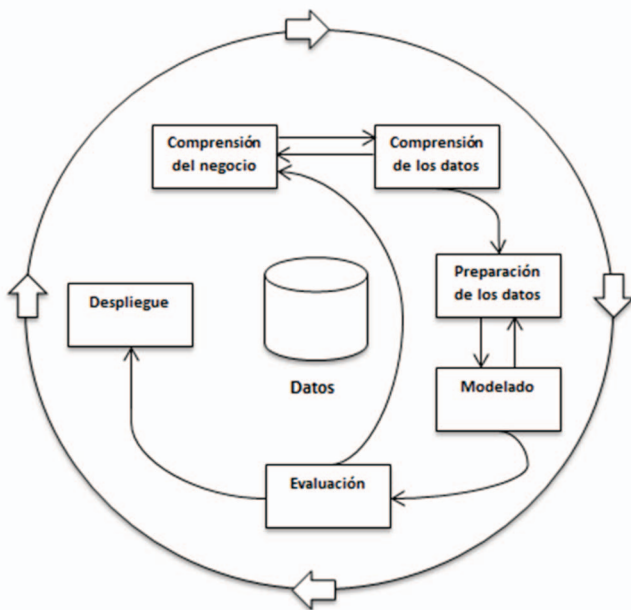


Figura 1. Fases del Modelo de Referencia CRISP-DM [7]

Para poder extraer conocimiento usando minería de datos se requiere la aplicación de uno o varios modelos o algoritmos. Estos modelos pueden ser descriptivos o predictivos [8]. Los modelos descriptivos usan modelos estadísticos como por ejemplo: distribución probabilística, correlación, regresión, análisis de clusters y análisis de discriminación. En los modelos descriptivos se conocen las características principales del conjunto de datos a analizar. Por otra parte, los modelos predictivos se basan mayormente en técnicas de aprendizaje automático (*machine learning*) y emplean inteligencia artificial [9]. Estos modelos se crean para predecir valores desconocidos de variables llamadas variables objetivo.

La variedad de aplicaciones para análisis de información en sectores comerciales, financieros, médicos, educativos o científicos, ha permitido el desarrollo de gran cantidad de métodos y algoritmos para predecir tendencias y descubrir patrones. Entre los más importantes se encuentran: *clustering*,

clasificación, regresión, análisis de desviación y análisis de cesta de mercado [10]. Estos métodos tienen una serie de soluciones como son: las redes neuronales, árboles de decisión o redes bayesianas. Por lo tanto es importante elegir el modelo de aprendizaje adecuado de acuerdo al problema a solucionar [11].

A. Minería de Datos Educativas

Al ser una disciplina nueva, no existe una definición ampliamente aceptada de MDE. Una definición que se ajusta a los objetivos de nuestra investigación es proporcionada por la International Society of Educational Data Mining “MDE es una disciplina en evolución, que tiene que ver con el desarrollo de métodos para explorar los tipos únicos de datos que provienen de ambientes educativos y por medio de la aplicación y uso de estos métodos una mejor comprensión de los estudiantes y el entorno en el que aprenden” [12].

La MDE usa métodos, herramientas y algoritmos de MD para investigar datos de interacciones de estudiantes y docentes con el sistema educativo, colaboración entre estudiantes, datos administrativos y datos demográficos. El proceso de MDE se muestra en la Figura 2. Se puede observar que existe una etapa previa de pre-procesamiento sobre los datos de los sistemas educativos. Los datos procesados son analizados eventualmente usando un repositorio basado en un modelo multidimensional [13] o relacional con algún algoritmo de MD y con la definición de un modelo. La interpretación del experimento permite evaluar y de ser necesario refinar el proceso con los resultados obtenidos.

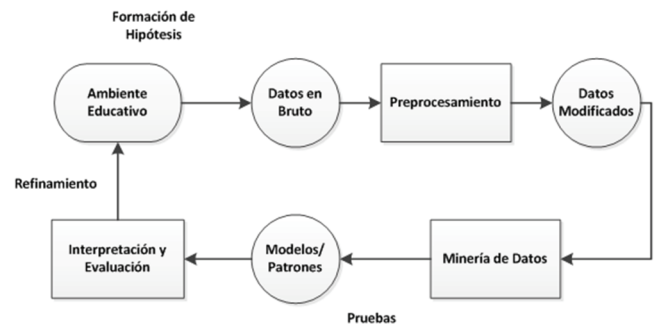


Figura 2. Descubrimiento de Conocimiento Educativo con Minería de Datos Basado en: [14]

La MDE explora el contexto organizacional de forma que puede ser clave para la mejora de indicadores como abandono estudiantil, tasa de graduación o en procesos de reestructuración y de gestión organizacional [15].

Existe una rama complementaria de MDE que se conoce como analítica del aprendizaje (AA), conocida en inglés como *learning analytics*. La principal diferencia entre ambas disciplinas es que MDE trata de descubrir conocimiento por medio de tareas automatizadas apoyadas por el juicio humano mientras que AA prioriza el descubrimiento de conocimiento basado en el juicio humano y se apoya en herramientas de automatización. Las diferencias de estas dos disciplinas se resumen en la Tabla 1 [16].

Tabla 1 Diferencias entre MDE y analítica del aprendizaje

Criterio	Diferencias	
	Minería de Datos Educativos	Analítica del Aprendizaje
Descubrimiento	El descubrimiento automatizado es clave; apalancar el conocimiento humano es una herramienta para cumplir esta meta	Apalancar el juicio humano es la clave; el descubrimiento automatizado es una herramienta para cumplir esta meta
Reducción y Holismo	Mayor énfasis en reducir en componentes y analizar las relaciones entre ellos	Mayor énfasis en entender los sistemas como un todo, en su complejidad total
Orígenes	Orígenes en software educativo y en modelamiento de estudiantes, con una comunidad de predecir los resultados de cursos	Orígenes en la web semántica "currículo inteligente", predicción de salidas
Adaptación y Personalización	Mayor importancia en adaptación automatizada (ej.: con la computadora y sin la intervención humana)	Mayor importancia en informar a instructores y estudiantes
Técnicas y Métodos	Clasificación, <i>clustering</i> , modelamiento bayesiano, minería de relaciones, descubrimiento con modelos, visualización	Análisis de redes sociales, análisis de sentimientos, analítica de influencia, análisis de discursos, predecir el éxito del estudiante

B. Clasificación de métodos, algoritmos y herramientas de MDE

La MDE integra métodos, algoritmos y técnicas con los cuales se puede realizar diferentes experimentos y diseñar modelos para predecir u obtener patrones de datos educativos. El objetivo de este artículo es presentar una visión holística de la disciplina, por lo cual se ha estudiado cuáles son los métodos, algoritmos y herramientas más usados por los investigadores de este campo científico, así como definir una clasificación global de los mismos.

Una clasificación de acuerdo al tipo de método de MD utilizado es propuesta por G. Siemens y R. Baker [17] y la detallamos a continuación:

- **Métodos de Predicción:** La meta de la predicción es diseñar un modelo que permita inferir en algún aspecto de los datos basándose en combinaciones de otros aspectos de los datos, por ejemplo se puede recolectar información de los abandonos de los estudiantes y con estos datos poder predecir para tomar acciones correctivas y preventivas (tutorías) sobre los nuevos estudiantes. Existen tres tipos que incluiríamos en esta categoría: clasificación, regresión y estimación de conocimiento latente.
- **Algoritmos de Descubrimiento de Estructuras:** Se trata de encontrar una estructura para los datos sin una idea previa de lo que se debería encontrar. El investigador trata de identificar cuál es la estructura natural de los datos. Dentro de esta clasificación se encuentran:

clustering, análisis de factores, análisis de redes sociales y descubrimiento de estructuras de dominio.

- **Minería de Relaciones:** El objetivo es descubrir relaciones entre ciertas variables de un conjunto de datos. Dentro de esta categoría se puede mencionar: asociación, correlación, minería secuencial de patrones y minería causal de datos.
- **Descubrimiento con Modelos:** Los resultados del análisis de minería de datos se utilizan para otro análisis posterior. Normalmente se obtiene un modelo a través de métodos de predicción.

Otra clasificación es presentada en el estudio de S. A. Kumar y Dr. Vijayalakshmi.M.N [18]. Esta clasificación muestra métodos de MD y los contextos educativos donde se aplica. El detalle se muestra en la Tabla 2.

Tabla 2 Clasificación de Métodos y Aplicaciones de MDE

Nombre del Método	Aplicaciones en MDE
Predicción	Detectar el comportamiento de los estudiantes. Desarrollo de modelos de dominio. Predecir los resultados de los estudiantes
Clustering	Descubrimiento de nuevos patrones de comportamiento de los estudiantes. Investigar similitudes y diferencias entre escuelas
Minería de Relaciones	Descubrir asociaciones curriculares en cursos
Descubrimiento con Modelos	Descubrir relaciones entre comportamiento de los estudiantes con las características de los estudiantes o variables contextuales. Análisis de preguntas de investigación dentro de una variedad de contextos
Destilación de Datos para Juicio Humano Secuencial	Identificación de patrones en el aprendizaje de estudiantes, comportamiento o colaboración; Etiquetar datos para uso posterior en desarrollo de nuevos modelos

C. Beneficios de la MDE

Existe una serie de procesos, servicios e indicadores organizacionales que se pueden beneficiar del uso de MDE en instituciones educativas. Los posibles beneficios del uso de MDE son [5] [19] [20] :

- Caracterizar el comportamiento y los logros de los alumnos.
- Mejorar el proceso de enseñanza/aprendizaje (matriculación, tutorías, evaluación, graduación).
- Mejorar indicadores de gestión como: disminuir abandono de alumnos, aumentar tasa de graduación, tutorías personalizadas.
- Mejorar la infraestructura.
- Optimizar el uso de aulas y laboratorios.
- Mejorar la eficiencia organizacional.

D. Instituciones que han aplicado con éxito MDE

Existen muchas instituciones educativas que están analizando sus datos y obteniendo conocimiento para que estudiantes, docentes o administradores puedan ser más eficientes. En esta sección citaremos algunos ejemplos de lo que se está haciendo.

Paul Smith's College usa analítica del aprendizaje para incrementar la tasa de graduación de sus egresados; la Washington University se asoció con la compañía "Persistence Plus" para mejorar las evaluaciones de sus cursos en línea; la Open University usa sus datos históricos para mejorar las tasas de retención de alumnos [21]; la University of Georgia llevó a cabo un experimento usando técnicas analíticas para predecir la tasa de graduación y abandono estudiantil en un ambiente en línea [22]. En la Purdue University han usado MD para determinar que la evaluación en etapas tempranas y de forma frecuente permite cambiar los hábitos de los estudiantes con calificaciones bajo la media en cursos introductorios. El equipo de investigación ha desarrollado un sistema de alerta académica temprana para saber el desempeño de los estudiantes [23].

Los artículos analizados en esta investigación contemplan interacciones de estudiantes con el sistema educativo, colaboración entre estudiantes, datos administrativos y demográficos.

III. RESULTADOS DE LA INVESTIGACION

Existen un sinnúmero de modelos y algoritmos de minería de datos que se pueden aplicar en el contexto de MDE. En este artículo se presenta un análisis de aquellos modelos sugeridos por investigadores que pueden ser aplicados a distintos escenarios prácticos. Como fuente primaria de la investigación se condujo una búsqueda basada en Internet usando Google Académico. El enfoque fue crear consultas de búsqueda sobre un rango de tiempo desde el año 2005 al 2015, cubriendo palabras clave que puedan identificar artículos relevantes acerca de EDM. Las palabras claves utilizadas en la búsqueda fueron *Educational Data Mining*, *Data Mining* y *Learning Analytics*. Posteriormente se descargó los artículos relacionados al tema llegando a un nivel de hasta 10 páginas de búsqueda ya que en estas primeras páginas de resultados de búsqueda aparecen los artículos con más citas y por ende aquellos considerados más importantes y cuyo criterio de relevancia es que presenten revisiones o experimentos de MDE. El total de artículos revisados fue de 300. La siguiente tarea realizada fue la de descartar libros o artículos técnicos ya que el alcance de nuestra investigación abarca únicamente artículos científicos. Los artículos seleccionados para posterior análisis fueron 65.

A estos artículos restantes se los clasificó para analizar metas, tipo de algoritmo, método o técnica usada, tipos de herramientas tecnológicas usadas y aplicaciones. En esta sección se presentan los resultados de este análisis que esperamos permita tener una visión holística de la disciplina y de los problemas que se han analizado y posibles implicaciones de investigación para el futuro.

A. Métodos y algoritmos de MDE más usados

De los 65 artículos sobre MDE que se estudiaron, 23 de ellos eran artículos de estado del arte, descriptivos o *surveys*. Estos artículos se excluyen del presente análisis, ya que solo se tiene en cuenta la investigación primaria y no la secundaria. Los 42 artículos restantes sí mostraban experimentos realizados sobre datos educacionales y el detalle de los métodos, técnicas o algoritmos más usados se muestran en la Tabla 3.

Tabla 3 Métodos y algoritmos más usados en MDE en los años 2005-2015

Método o algoritmo	Nro. de experimentos
Clasificación	18
Clustering	14
Asociación	12
Predicción	8
Minería Secuencial Patrones	4
Outlier Detection	3
Minería de Texto	3
Análisis de Redes Sociales	2
Q-Matrix	2
Análisis de Factores	1
Procesos de Descisión Markov	1
LIFT	1

Se puede observar que la mayor cantidad de experimentos en MDE usan clasificación, clustering, asociación y predicción para resolver los casos de estudio. Es importante indicar que en algunos experimentos se usó más de una técnica por lo que el total no suma 42.

B. Herramientas de MDE más usadas

Como parte del estudio también se analizó los frameworks y las herramientas usadas para los experimentos realizados en los 42 artículos del estudio. El framework más usado para guiar el proceso de MDE es CRISP. Existe también artículos que trabajan con el enfoque PDCA (Plan, Do, Check, Act). Las herramientas usadas en los experimentos se muestran en la Tabla 4.

C. Procesos educativos donde se aplican los métodos y algoritmos

En esta sección se presenta un análisis realizado sobre la muestra de los 42 artículos de experimentos que fueron considerados en este estudio. De la revisión se observa que existen áreas prioritarias de experimentación como se muestra en la Tabla 5. Observamos que el escenario donde más experimentos se realizaron es el de predicción de evaluaciones o rendimiento de los estudiantes (21 experimentos), además 6 experimentos se realizaron para predecir el perfil y comportamiento de aprendizaje de los estudiantes. Se observa también que muy pocos investigadores analizan el potencial de dirigir experimentos para mejorar la toma de decisiones por medio de predecir los procesos organizacionales que son sujetos a perfeccionamiento.

Tabla 4 Herramientas usadas en experimentos de MDE

Herramienta	Meta
Cold Fusion	Lenguaje de programación para desarrollo de aplicaciones web diseñado para conectar de manera fácil páginas HTML a una base de datos
Microsoft SQL Server Analysis Services	Ofrece funciones de procesamiento analítico en línea (OLAP) y MD para aplicaciones de <i>Business Intelligence</i>
EDM workbench	Herramienta para descubrir información de datos de registro
Inq-ITS System	Realiza el seguimiento del progreso de los estudiantes usando algoritmos de minería de datos y reportes
Moodle Datamining Tool	Herramienta de minería de datos para Moodle desarrollado en Java, basado en Keel
RapidMiner	Plataforma libre para análisis predictivo usando machine learning y técnicas analíticas
R	Lenguaje y entorno de programación para análisis estadístico y gráfico
Weka	Colección de algoritmos de aprendizaje automático para tareas de minería de datos
KEEL	Herramienta de software usado en diferentes tareas de descubrimiento de conocimiento
SNAPP	Permite evaluar patrones de comportamiento del estudiante con el diseño de actividades de aprendizaje
AHA System	Recomendar los mejores links que puede visitar en un siguiente paso el estudiante
Knime	Plataforma de analítica para tratar de descubrir patrones en los datos y predecir tendencias
GISMO	Monitoreo gráfico interactivo que provee visualización de las actividades de los estudiantes
SPSS	Predecir de forma confiable el futuro para una mejor toma de decisiones
Sequential Mining Tool	Minería que ayuda a los profesores a descubrir información importante de alumnos
KEA	Software que usa un algoritmo para extracción de palabras claves de documentos de texto
DB Miner	Minería para descubrir distintos niveles de conocimiento en grandes bases de datos relacionales

En la investigación de N. Delavari, S. Phon-amnuaisuk y M. Beikzadeh [24] se presenta una clasificación de los procesos en instituciones de educación superior que podrían ser susceptibles de mejoras a través de técnicas de MD. Entre los procesos mencionados en el estudio se encuentran: evaluación, planificación, matriculación, bienestar del estudiante, tutorías, rendimiento estudiantil y docente. Podemos concluir que existe mucho campo de experimentación en las áreas de planificación y matriculación, ya que en los artículos escritos en los años objeto de este estudio existe poco análisis en estos procesos.

Tabla 5 Escenarios de utilidad de MDE

Escenarios de utilidad de MDE	Total
Predecir evaluaciones finales o rendimiento de los estudiantes	21
Predecir el perfil y comportamiento de aprendizaje de los estudiantes	6
Mejorar el soporte del docente	3
Mejorar la gestión de la colaboración en ambientes educacionales	3
Generación de consejos basados en datos históricos de estudiantes	2
Recomendar estudiantes links de navegación en sistemas web	2
Descubrir representaciones alternativas de datos en foros	1
Predecir abandono estudiantil	1
Selección de prácticas de forma adaptativa	1
Predecir como diferencias entre entornos de aprendizaje pueden aumentar o disminuir la probabilidad de juego.	1
Mejorar los procesos de toma de decisiones	1

IV. CONCLUSIONES

Este artículo presenta una visión holística de MDE y busca ser una buena fuente para investigadores que deseen experimentar con MD en el campo educativo en áreas como evaluación, matriculación, planificación, bienestar estudiantil, marketing, etc. MDE se proyecta como una disciplina esencial para la gestión universitaria que otorga visibilidad a los directivos para mejorar la toma de decisiones.

La MDE junto con la AA ha realizado ya contribuciones importantes en instituciones educativas para la mejora de los procesos del estudiante y el docente. En este artículo se han detallado los algoritmos y métodos existentes de MD que han sido usados en experimentos de MDE. Los métodos más utilizados son: clasificación, *clustering* y asociación. Además, hemos presentado las herramientas que soportan el análisis de MD para cubrir distintos escenarios y algoritmos.

De la investigación se observa que la mayoría de los experimentos realizados están enfocados en los procesos de enseñanza/aprendizaje, como predicción de evaluaciones, rendimiento y perfil de los estudiantes, así como proveer retroalimentación a los tutores o docentes.

Muchas áreas de interés para estudios posteriores han sido detectadas en este artículo. Áreas con gran potencial son la de planificación educativa y bienestar estudiantil de las instituciones de educación superior. Además se sugiere en investigaciones futuras profundizar en temas como: analítica del aprendizaje social, analítica del aprendizaje multimodal o MDE con Big Data.

REFERENCIAS BIBLIOGRÁFICAS

- [1] W. Ian and F. Eibe, *Data Mining: Practical Machine Learning Tools and Techniques*, Second Edi. San Francisco: Elsevier B.V., 2005.
- [2] R. Jindal and M. D. Borah, "A Survey on Educational Data Mining and Research Trends," *Int. J. Database Manag. Syst.*, vol. 5, no. 3, pp. 53–73, 2013.
- [3] R. A. Huebner, "A Survey of Educational Data-Mining Research," *Res. High. Educ. J.*, pp. 1–13, 2013.
- [4] S. Mitra, S. K. Pal, and P. Mitra, "Data mining in soft computing framework: A survey," *IEEE Trans. Neural Networks*, vol. 13, no. 1, pp. 3–14, 2002.
- [5] L. A. O. P. Xavier, M. Oliveira, and E. K. Teixeira, "Teorias utilizadas nas investigações sobre gestão do conhecimento," *RISTI - Rev. Iber. Sist. e Tecnol. Inf.*, vol. 10, no. 1, pp. 1–18, 2012.
- [6] Y. Peng, G. Kou, Y. Shi, and Z. Chen, "a Descriptive Framework for the Field of Data Mining and Knowledge Discovery," *Int. J. Inf. Technol. Decis. Mak.*, vol. 07, no. 04, pp. 639–682, 2008.
- [7] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, and W. Rudiger, "Crisp-Dm 1.0," *Cris. Consort.*, p. 76, 2000.
- [8] L.-D. Chen, T. Sakaguchi, and M. N. Frolick, "Data Mining Methods, Applications, and Tools," *Inf. Syst. Manag.*, vol. 17, no. 1, pp. 65–70, Jan. 2000.
- [9] J. L. Hung and K. Zhang, "Data Mining Applications to Online Learning," *World Conf. E-Learning Corp. Gov. Heal. High. Educ.*, pp. 2014–2021, 2006.
- [10] M. J. Jafar, "A Tools-Based Approach to Teaching Data Mining Methods," *J. Inf. Technol. Educ.*, vol. 9, pp. 1–29, 2010.
- [11] a F. Elgamal, "An Educational Data Mining Model for Predicting Student Performance in Programming Course," *Int. J. Comput. Appl.*, vol. 70, no. 17, pp. 22–28, 2013.
- [12] International Society Educational Data Mining, "Educational Data Mining," 2015. [Online]. Available: <http://www.educationaldatamining.org/>. [Accessed: 29-Jun-2015].
- [13] S. Luján-Mora, J. Trujillo, and I. Y. Song, "Extending the UML for Multidimensional Modeling," in *Lecture Notes in Computer Science*, 2002, vol. 2460, pp. 290–304.
- [14] S. Ventura and C. Romero, "Data mining in education," *WIREs Data Min. Know Discov*, no. August, pp. 12–27, 2013.
- [15] M. Bienkowski, M. Feng, and B. Means, "Enhancing teaching and learning through educational data mining and learning analytics: An issue brief," *Dep. Educ. Off. Educ. Technol.*, pp. 1–57, 2012.
- [16] G. Siemens and R. Baker, "Learning analytics and educational data mining: towards communication and collaboration," *Proc. 2nd Int. Conf. Learn. Anal. Knowl.*, pp. 252 – 254, 2012.
- [17] R. Baker and G. Siemens, *Educational Data Mining and Learning Analytics*. Cambridge, UK, 2011.
- [18] S. A. Kumar and Dr. Vijayalakshmi.M.N, "A Novel Approach in Data Mining Techniques for Educational Data S.Anupama," *3rd Int. Conf. Mach. Learn. Comput. (ICMLC 2011) A*, no. Icmcl, pp. 152–154, 2011.
- [19] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1432–1462, 2014.
- [20] R. Baker and K. Yacef, "The state of educational data mining in 2009: A review and future visions," *J. Educ. Data Min.*, vol. 1, no. 1, pp. 3–17, 2009.
- [21] J. Bichsel, "Analytics in Higher Education Benefits, Barriers, Progress and Recommendations," 2012. [Online]. Available: <https://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf>.
- [22] L. Morris, S.-S. Wu, and C. Finnegan, "Predicting retention in online general education courses," *Am. J. Distance Educ.*, vol. 19, no. 1, pp. 23–36, 2015.
- [23] P. Baepler and C. J. Murdoch, "Academic Analytics and Data Mining in Higher Education," *Int. J. Scholarsh. Teach. Learn.*, vol. 4, no. 2, pp. 1–9, 2010.
- [24] N. Delavari, S. Phon-amnuaisuk, and M. Beikzadeh, "Data Mining Application in Higher Learning Institutions," *Informatics Educ.*, vol. 7, pp. 31–54, 2008.