
DISEÑO DE UN SISTEMA DE INTELIGENCIA DE NEGOCIOS PARA CONTRIBUIR A LA PERMANENCIA DE LOS ESTUDIANTES EN LA FUNDACIÓN UNIVERSITARIA DE POPAYÁN

Zuleidy Yaruro Orozco, Maycol Pérez

RESUMEN

Actualmente la inteligencia de negocios se ha convertido en una importante forma de implementar estrategias que ayudan a las diferentes entidades a predecir situaciones que puedan presentarse con el fin de poder tomar decisiones.

En la Fundación Universitaria de Popayán hay un problema de abandono por parte de los estudiantes que cada vez va en aumento. Es por esto, que a continuación, se propone diseñar y desarrollar un sistema de inteligencia de negocios, con el cual, a partir de los resultados arrojados se espera tomar decisiones de forma inmediata para disminuir la deserción de los estudiantes de la institución, y de esta manera aumentar la tasa de permanencia de estas personas.

El sistema propuesto es un tablero BI de Google Data Studio en donde se logró evidenciar que hay factores, como, por ejemplo, la edad y el promedio de calificaciones que efectivamente influyen en la decisión de un estudiante de abandonar o permanecer en una institución de educación superior.

Palabras clave: negocios, inteligencia, estudiantes, deserción, universidad, predicción

1. INTRODUCCIÓN

Una de las problemáticas más relevantes que presenta la educación superior en Colombia tiene que ver con las altas tasas de deserción académica. El número de estudiantes que logran culminar su carrera disminuye con el paso del tiempo y se presenta por lo general en los primeros cinco semestres [1].

El informe sobre educación superior en América Latina y el Caribe, presentado por la UNESCO [2], estudia la deserción en esta área geográfica, promediando las áreas de conocimiento con las tasas de deserción de los países de estas regiones, teniendo como resultado que la tecnología y la ingeniería tienen una tasa de graduación del 38,5 % [1].

Hay instituciones de educación superior que no logran realizar un seguimiento continuo de sus estudiantes debido a la falta de organización de los datos [3]; teniendo en cuenta que, al no tener un correcto orden de esta información, no se puede implementar algún tipo de plan para evitar que los estudiantes abandonen la institución. Sin embargo, algunas instituciones educativas de educación superior vienen adelantando esfuerzos en lograr ordenar estos datos para poder tener la trazabilidad necesaria y de esta forma poder implementar mejoras con el fin de evitar el abandono de los estudiantes, pero no se han logrado avances significativos en el campo.

Existen sistemas de inteligencia de negocios en distintas universidades, muchos de estos se centran en mostrar estadísticas [4], pero, no predicen o no se relacionan directamente con el abandono estudiantil ni tampoco indican las variables que afectan la decisión por el cual un estudiante decide dejar a medias sus estudios profesionales.

Por otra parte, hay trabajos que estudian el abandono de los estudiantes en instituciones de educación superior [5]. Estas investigaciones son llevadas a cabo por medio de datasets de entrenamientos, implementación de distintos algoritmos e inteligencia artificial, pero, no se visualizan los datos de una forma amigable, fácil de entender para poder tomar decisiones que ayuden a disminuir el abandono por parte de los estudiantes de estas universidades.

Una de las instituciones de educación superior del departamento del Cauca que ha venido adelantando procesos de organización de sus datos es la Fundación Universitaria de Popayán, ahora bien, aquí no se cuenta con un sistema de visualización de datos que sea amigable y fácil de entender como lo es, por ejemplo, un sistema de inteligencia de negocios con el que se pueda tomar decisiones a partir de la información que este pueda brindar.

Teniendo en cuenta lo anterior, se planteó como objetivo principal el de diseñar un sistema de inteligencia de negocios para hacer un seguimiento continuo y así favorecer a la permanencia estudiantil de los estudiantes de la Fundación Universitaria de Popayán.

2. MÉTODOS

La inteligencia de negocios (Business Intelligence o BI) es un conjunto de herramientas y metodologías que permiten el análisis, tratamiento, exploración y transformación de los datos con el fin de mejorar en gran medida la toma de decisiones en una compañía o entidad [4].

A. Metodología

Para el desarrollo del presente artículo se trabajó con varias metodologías las cuales tuvieron su aporte en diferentes aspectos. Las metodologías utilizadas son las siguientes:

Metodología Mixta: Esta metodología de investigación consiste principalmente en la recopilación, análisis de los datos de tipos cuantitativos y cualitativos [6].

Metodología Transversal: Es una forma de investigación de tipo estadístico que se centra en determinar la condición de una población en especial, para el caso del presente trabajo, se focaliza en el análisis de los datos de los estudiantes de la FUP con el fin de determinar posiblemente quién abandona la carrera profesional y quién no[7].

Metodología CRISP-DM: Por su siglas en inglés Cross-Industry Standard Process for Data Mining(Proceso estándar entre industrias para la minería de datos) es una metodología que tendrá su aporte principalmente en la parte del modelado

de los datos en donde se podrá aplicar el ciclo de vida de un proyecto de análisis de datos [8].

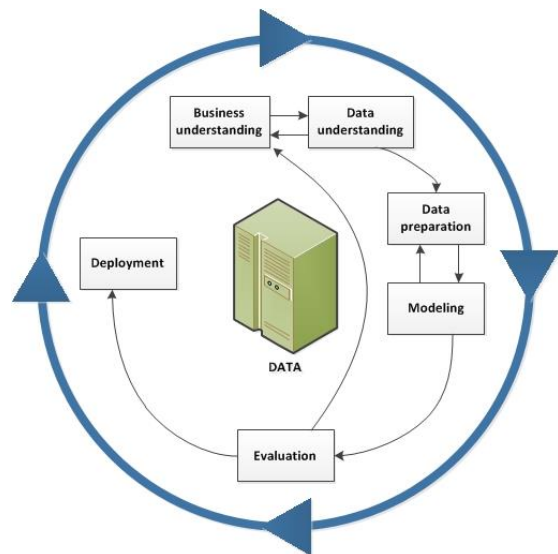


Figura 1. Ciclo Metodología CRISP DM [8].

La metodología CRISP-DM abarca cada aspecto de un proyecto de análisis de datos por medio de sus seis fases [8]:

Business Understanding (Comprensión del Negocio): es la fase inicial del ciclo de vida de la CRISP-DM, la cual está enfocada principalmente en entender cómo funciona el negocio [8]. En este punto es importante destacar que se realizó un análisis previo sobre la FUP en cuanto al manejo de herramientas BI para análisis de los datos de sus estudiantes, evidenciando que la institución no se encuentra a la vanguardia de estas herramientas.

Data Understanding (Comprensión de los datos): es la fase de reconocimiento de los datos en donde se podrán identificar problemas de calidad [8]. En cuanto a los datos del presente proyecto se trabajó con un dataset proporcionado por la FUP en el que había diferentes variables con las cuales se realizó el respectivo sistema BI, como, por ejemplo, estado de matrícula del estudiante, promedio de las diferentes áreas, si el estudiante es o no desertor, entre otras.

Data Prepatation (Preparación de Datos): fase en la cual se realizan las actividades que den lugar a la construcción del conjunto final de datos [8]. Precisamente el dataset proporcionado por la FUP se tuvo que someter a una previa preparación, allí

se pudo eliminar datos a los cuales les faltaban campos por llenar; también los promedio que tenían números flotantes muy grandes se convirtieron a números con máximo dos decimales.

Modeling (Modelado): en esta etapa se aplican las técnicas pertinentes de modelamiento de datos [8]. Aquí fue necesario abstraer algunos puntos importantes del problema como, por ejemplo, la definición de los requerimientos del negocio y dimensiones.

Requerimientos del modelo de negocio: la definición de los requerimientos se centra en los usuarios de la institución los cuales serán los encargados de realizar el respectivo análisis con el propósito de tomar decisiones; el sistema presentado a continuación será un modelo que la institución podrá tomar más adelante con datos más precisos y actualizados. Algunos de los requerimientos más importantes son:

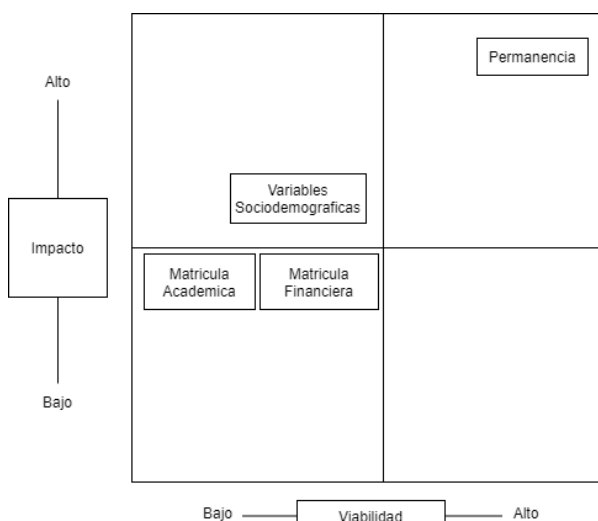


Figura 2. Priorización Requerimientos del Negocio

Permanencia: En este punto se tratará lo relacionado a los estudiantes que desertan y los que no lo hacen.

Variables Sociodemográficas: En este tema se manejan algunas variables importantes cómo lo son los promedios y las edades de los estudiantes.

Matricula Académica y Matricula Financiera: juegan un papel muy importante dentro de la priorización de los requerimientos del sistema ya que si un estudiante abandona no es posible legalizar matrícula para esta persona.

Dimensiones				
Tema Negocio	Promedio Carrera	Promedio lcfes	Edad	Desertor
Permanencia	x			x
Variables Sociodemográficas	x		x	x
Matricula Financiera	x			
Matricula Académica	x			

Tabla 1. Dimensiones del Negocio

Evaluation (Evaluación): en esta fase de la metodología es importante evaluar el modelo de los datos antes de realizar el despliegue[8]. Por lo que se decide implementar las Heurísticas de Nielsen, las cuales son un conjunto de técnicas para probar las interfaces de los desarrollos realizados por diseñadores y desarrolladores [9]; también se utilizó SUS (System Usability Scale, Sistema de Escalas de Usabilidad) que es un método rápido para evaluar la usabilidad de los sistemas [10], este último fue implementado con algunos estudiantes del programa de ingeniería de sistemas en la FUP, arrojando un porcentaje mayor al 88%, lo cual indica que la satisfacción del usuario superó las expectativas.

Deployment (Despliegue): esta fase cubre la parte del alojamiento del sistema y posterior presentación al cliente [8].

Mapeo sistemático: Es una metodología que se centra en la recopilación, clasificación y revisión de los artículos que servirán de soporte para el presente trabajo [11].

Según la metodología “Mapeo Sistemático”, se definieron las cadenas de búsquedas en primera instancia.

- (student) AND (prediction) AND (university OR college) AND (desertion OR abandonment OR dropout) AND (data mining) AND (machine learning)
- (intelligence) AND (business) AND (desertion OR dropout) AND (student) AND (prediction) AND (education)

En segundo lugar, se procedió a utilizar las cadenas de búsquedas en cada de unas de las respectivas bases de datos con el fin de encontrar los artículos que tuvieran similitudes al presente trabajo. Los trabajos de investigación fueron

tomados en cuenta de acuerdo con su título y año de publicación.

Los motores de búsquedas utilizados son los siguientes:

- IEEE Explore
- Science Direct
- Google Scholar

	IEEE	Science Direct	Google Scholar	Total
(student) AND (prediction) AND (university OR college) AND (desertion OR abandonment OR dropout) AND (data mining) AND (machine learning)	15	5	5	25
(intelligence) AND (business) AND (desertion OR dropout) AND (student OR) AND (prediction) AND (education)	0	10	10	20

Tabla 2. Resultados Bibliográficos

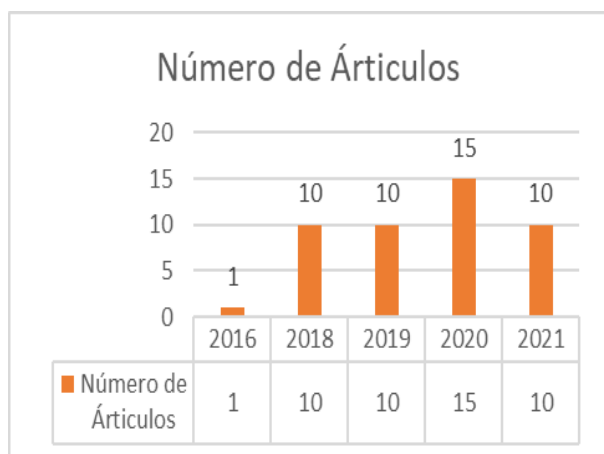


Figura 3. Número de Artículos por Año

Posteriormente, se realizó la selección de los artículos que tuvieran un alto grado de interés; para este proceso se tuvo en cuenta varios aspectos importantes, como, por ejemplo, que el título estuviese relacionado con el tema de predicción de abandono estudiantil en instituciones de educación superior e inteligencia de negocios, el año de publicación y contenido del artículo referente al tema del presente informe.

B. Estado del Arte

Es muy importante realizar una retrospectiva de los trabajos encontrados en las diferentes bases de datos consultadas, los cuales están relacionados y

que tienen mayor relevancia para el presente proyecto, con el objetivo de conocer las tecnologías utilizadas que puedan ser de utilidad y así logren aportar a la solución del problema de la deserción estudiantil en la Fundación Universitaria de Popayán. Los trabajos descritos a continuación están ordenados por trascendencia en el tiempo.

En 2017 Aparna Gopalakrishnan et al. en Estados Unidos, propusieron un sistema de análisis de datos en el que utilizaban diferentes métodos de minería de datos para evitar la deserción de estudiantes universitarios y de esta manera aumentar la tasa de graduados [12]. Algunos de los métodos utilizados fueron los algoritmos de clasificación. Los resultados obtenidos de esta investigación permitieron tomar acciones inmediatas para proporcionar ayudas a los estudiantes los cuales el sistema arrojaba como posibles desertores. De todas formas, este trabajo de investigación no hizo ningún tipo de relación con la inteligencia de negocios.

Por su parte, en 2018 Y. M. Pérez-Pérez et al. en la ciudad de Ocaña, Colombia, desarrollaron un sistema de inteligencia de negocio aplicado a instituciones de educación superior enfocándose principalmente en procesos de gestión administrativa [4]. Entre los resultados se encontraron coincidencias en cuanto al sistema desarrollado y el proceso en el que se implementó, por lo que resultó muy útil. La lectura de este artículo reveló que este sistema de inteligencia de negocios no proporciona predicción sobre la deserción de los estudiantes.

En 2019 Rodrigo da Fonseca Silveira et al. en Brasilia, Brasil, realizaron un trabajo de investigación en el que implementaron un análisis de datos sobre el abandono de los estudiantes de pregrado de algunas ingenierías de la universidad de Brasilia donde utilizan algunas técnicas de minería de datos [6]. Sin embargo, se evidenció que no utilizan la inteligencia de negocios para mostrar sus datos de forma más amigable.

También, en 2021 Warit Tenpipat et al. en Tailandia, desarrollaron un trabajo de investigación en el que estudiaron los factores causantes de la deserción de estudiantes universitarios por medio de la creación de modelos de clasificación con el fin de predecir si un alumno será posiblemente un desertor o no [13]. En cuanto a los resultados, estos muestran que la precisión de la predicción

ronda el 92%. Se pudo constatar que este trabajo de investigación tampoco utiliza la inteligencia de negocios para mostrar con mayor agrado sus resultados.

Teniendo en cuenta los trabajos descritos anteriormente, se evidencia que definitivamente no hay trabajos de investigación que estudien la deserción de los estudiantes desde el punto de vista de la inteligencia de negocios; es por esto, que se decide realizar el estudio sobre aquellos estudiantes que posiblemente abandonarían la institución por medio de un sistema de inteligencia de negocios.

C. ARQUITECTURA

Para el diseño e implementación del sistema de inteligencia de negocios se utilizó la arquitectura Kimball.

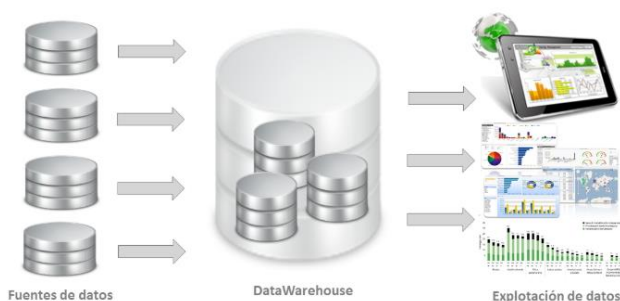


Figura 4. Arquitectura Kimball [14]

La arquitectura Kimball es un conjunto de metodologías empleada para la construcción de sistemas de inteligencia de negocios, que ayuda a la toma de decisiones de una empresa. Esta arquitectura tiene tres enfoques principales [14]:

- Fuentes de datos: hace referencia al conjunto de datos que se utilizaría para realizar el sistema de inteligencia de negocios [14]. Para este punto clave se utilizó un dataset proporcionado por la FUP en donde se pueden encontrar variables, como, por ejemplo, promedios de materias, promedio icfes, si es desortor.
- ETL: Por sus siglas en inglés “Extract, Transform, and Load”, es un término que hace referencia a la integración de datos con el fin de construir un almacén con esta información [14]. En este paso se realizó el

ETL de la fuente de los datos proporcionados por la universidad, del cual se pudo extraer los datos para realizar el respectivo tablero.

- Exploración de Datos: en este punto se logró conectar el ETL a la herramienta BI DataStudio de Google en donde se lograron visualizaciones de los resultados muy acertados para la toma de decisiones.

3. RESULTADOS



Figura 5. Tablero BI

De acuerdo con el tablero BI elaborado se llegó a las siguientes conclusiones:

- El 46,5% de los estudiantes que se analizaron en el sistema de inteligencia de negocios desertaron, el otro 53.5% permanecieron en la universidad.
- Los estudiantes con mejores promedios se encuentran por encima de los 25 años.
- Los estudiantes que abandonaron la universidad tienen un promedio general menos favorable que los que no abandonan, esto significa que, los alumnos posiblemente se sientan desmotivados debido a su desempeño académico.
- Los estudiantes que abandonaron la universidad tienen en promedio entre 25 y 30 años.

Estos resultados permitirán a las directivas tomar acciones oportunas con el fin de evitar que los estudiantes abandonen la universidad y así aumentar en gran medida la permanencia de estas personas en la universidad.

4. DISCUSIÓN (O ANÁLISIS DE RESULTADOS)

- A lo largo del desarrollo del presente proyecto se pudo evidenciar que los datos que proporcionó la universidad se consideran no viables, debido a que realizar un sistema de BI con tan pocas variables se vuelve muy limitado; es por esto que se cree que hizo falta adicionar algunas columnas más, como, por ejemplo, fechas de matrículas, históricos, valores sociodemográficos como lo son el género y la edad. Sin embargo, de la literatura se pudo obtener la información básica de cómo se podría obtener el modelo del tablero que se desarrolló, el cual tiene un diseño preliminar que sirve de referencia para ser aplicado con una fuente de datos más confiable.
- El modelado de datos argumentado anteriormente facilitará la elaboración del sistema BI para toma de decisiones de la institución siempre y cuando se tengan en cuenta las variables que se proponen en el presente informe.

5. CONCLUSIÓN

- Durante el diseño y desarrollo del tablero BI se pudo apreciar cómo la inteligencia de negocios logra jugar un papel muy importante a la hora de realizar análisis de datos con el fin de tomar decisiones acertadas, ya que logra notablemente la interactividad entre los usuarios y los resultados agradables, haciendo que las herramientas de BI cada vez más sean utilizadas para lograr mejores beneficios sobre la entidad que las usa.
- De igual forma, se evidenció que la inteligencia de negocios ayuda a visualizar y controlar lo que sucede en cada línea del negocio, por la medición de métricas e indicadores específicos, haciendo así que se actué con seguridad en la toma de decisiones soportada en información confiable y real.

6. REFERENCIAS BIBLIOGRÁFICAS

- [1] C. Guzman Ruiz, D. M. Duran Muriel, J. Franco Gallego, E. Castaño Velez, S. Gallon Gomez, y K. Gomez Portilla, *Desercion estudiantil en la educacion superior colombiana: metodologia de seguimiento, diagnostico y elementos para su prevencion*. Bogot? Ministerio de Educacion Nacional, 2009.
- [2] H. Grimaldo *et al.*, «COORDINADOR TEMÁTICO», p. 204.
- [3] T. Devasia, Vinushree T P, y V. Hegde, «Prediction of students performance using Educational Data Mining», en *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*, Ernakulam, India, mar. 2016, pp. 91-95. doi: 10.1109/SAPIENCE.2016.7684167.
- [4] Y. M. Pérez-Pérez, A. A. Rosado-Gómez, y A. M. Puentes-Velásquez, «Application of business intelligence in the quality management of higher education institutions», *J. Phys. Conf. Ser.*, vol. 1126, p. 012053, nov. 2018, doi: 10.1088/1742-6596/1126/1/012053.
- [5] H. A. Combita Niño, J. P. Cómbita Niño, y R. Morales Ortega, «Business intelligence governance framework in a university: Universidad de la costa case study», *Int. J. Inf. Manag.*, vol. 50, pp. 405-412, feb. 2020, doi: 10.1016/j.ijinfomgt.2018.11.012.
- [6] R. da Fonseca Silveira, M. Holanda, M. de Carvalho Victorino, y M. Ladeira, «Educational Data Mining: Analysis of Drop out of Engineering Majors at the UnB - Brazil», en *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, Boca Raton, FL, USA, dic. 2019, pp. 259-262. doi: 10.1109/ICMLA.2019.00048.
- [7] E. J. Chacón Guerrero y R. A. Eslava Zapata, «Aplicaciones de Software Científico para el análisis de datos en diseños mixtos de

investigación.», *Eco Matemático*, vol. 8, n.º 1, pp. 106-115, ene. 2017, doi: 10.22463/17948231.1481.

[8] J. C. M. Zapata y N. Gil, «Incorporation of both pre-conceptual schemas and goal diagrams in CRISP-DM», en *2011 6th Colombian Computing Congress (CCC)*, Manizales, Colombia, may 2011, pp. 1-6. doi: 10.1109/COLOMCC.2011.5936284.

[9] M. Andreu-Vall y M.-C. Marcos, «Evaluación de sitios web multilingües: metodología y herramienta heurística», *El Prof. Inf.*, vol. 21, n.º 3, pp. 254-260, may 2012, doi: 10.3145/epi.2012.may.05.

[10] E. Universitat Politècnica de València, «Universitat Politècnica de València», *Ing. Agua*, vol. 18, n.º 1, p. ix, sep. 2014, doi: 10.4995/ia.2014.3293.

[11] W. A. Ortega-Ordóñez, C. J. Pardo-Calvache, y F. J. Pino-Correa, «Systematic mapping on the agility evaluation in software development organizations», *ITECKNE*, vol. 16, n.º 1, pp. 64-76, jun. 2019, doi: 10.15332/iteckne.v16i1.2162.

[12] A. Gopalakrishnan, R. Kased, H. Yang, M. B. Love, C. Graterol, y A. Shada, «A multifaceted data mining approach to understanding what factors lead college students to persist and graduate», en *2017 Computing Conference*, London, jul. 2017, pp. 372-381. doi: 10.1109/SAI.2017.8252128.

[13] W. Tenpipat y K. Akkarajitsakul, «Student Dropout Prediction: A KMUTT Case Study», en *2020 1st International Conference on Big Data Analytics and Practices (IBDAP)*, Bangkok, Thailand, sep. 2020, pp. 1-5. doi: 10.1109/IBDAP50342.2020.9245457.

[14] E. M. Rojas, «Implementación de Inteligencia de Negocios en proyectos utilizando Scrum y arquitectura Kimball Implementation of Business Intelligence in projects using Scrum and Kimball architecture», p. 19, 2019.