



TRABAJO DE GRADO
Opción Seminario-Diplomado.

**ALGORITMO COMPUTACIONAL PARA EL ANÁLISIS Y TOMA DE DECISIONES
EN DATOS DE DESERCIÓN ESCOLAR, UTILIZANDO ESTRATEGIAS DE
MACHINE LEARNING**

Corporación Universitaria Remington.
Facultad de ingenierías
Seminario de Machine Learning

Estudiante:
JOSE CAMILO CASTAÑO DIAZ
Tutor: Juan Carlos Briñez de León
Opción de Trabajo de grado Seminario-Diplomado.
2024.

Tabla de Contenidos

Resumen	3
Palabras clave	4
Marco conceptual y contextual	5
Pregunta problema	10
Acercamiento a los datos	12
Posibles aplicaciones	13
Aproximaciones con graficos	14
Gráfica 1	14
Grafica 2	15
Grafica 3	16
Grafica 4	17
Grafica 5	18
Grafica 6	19
Objetivos	20
Objetivos generales	20
Objetivos específicos	20
Desarrollo e implementación del aprendizaje	20
Procesamiento de datos	21
Modelo de toma de decisiones	22
Implementación en contextos reales	23
Resultados adicionales	26
Figuras y tablas	27
Conclusiones	29
Referencias	30

Resumen

El abandono de los estudios antes de culminar el ciclo educativo representa un desafío multidimensional con repercusiones personales y sociales significativas. A través de este proyecto, se examinaron de manera exhaustiva los factores determinantes, patrones y estrategias para abordar este fenómeno en el contexto colombiano.

Los hallazgos revelaron una tendencia decreciente en las tasas nacionales de deserción durante la última década, aunque persisten brechas preocupantes en ciertos segmentos poblacionales y regiones geográficas. Las disparidades más notorias se evidenciaron entre las zonas urbanas y rurales, con índices de abandono considerablemente más elevados en estas últimas, reflejando las barreras adicionales que enfrentan las comunidades alejadas de los centros urbanos.

Asimismo, los análisis indicaron que los niveles superiores de educación, particularmente la enseñanza media, exhiben mayores tasas de deserción en comparación con la primaria, destacando la urgencia de implementar estrategias de retención específicas en estas etapas cruciales de formación académica.

Palabras clave

Deserción escolar, Datos, Factores de riesgo educativos, Cobertura educativa, Calidad educativa, Machine learning.

Marco conceptual y contextual

En la era contemporánea, las instituciones educativas han emprendido una transformación digital que les ha permitido recopilar y sistematizar información valiosa sobre la problemática de la deserción escolar. Estos registros sistemáticos no solo contienen datos sobre las tasas de abandono estudiantil, sino también sobre las decisiones tomadas por los planteles, las estrategias implementadas y los factores contextuales asociados a cada caso en particular.

Este acervo de datos digitalizados constituye un activo clave para comprender de manera integral la confluencia de factores multidimensionales que inciden en la deserción. Mediante análisis profundos, es posible identificar patrones, tendencias y relaciones subyacentes entre aspectos como las características socioeconómicas del alumnado, los entornos familiares, la calidad de la enseñanza impartida, los recursos institucionales disponibles, las políticas educativas vigentes y una multiplicidad de condicionantes internos y externos.

Al procesar estos grandes volúmenes de información histórica y en tiempo real, se pueden desarrollar modelos predictivos y sistemas de alertas tempranas que permitan anticipar situaciones de riesgo y activar medidas preventivas oportunas. Además, esta inteligencia de datos facilita la evaluación objetiva del impacto de las estrategias contra la deserción puestas en marcha, posibilitando su mejora continua y adaptación al contexto.

No obstante, la sola recopilación de datos no es suficiente. Es imperativo analizar e interpretar adecuadamente esta información aplicando metodologías rigurosas y un enfoque colaborativo que involucre a todos los actores clave: autoridades educativas, instituciones, docentes, estudiantes y sus familias. Solo entonces, los hallazgos obtenidos podrán traducirse en políticas, programas y acciones efectivas para prevenir y mitigar el abandono escolar en todas sus manifestaciones.

Esta visión resalta el valor estratégico que representan los datos digitalizados para abordar un reto tan complejo como la deserción estudiantil. Al integrar esta perspectiva con los aspectos previamente mencionados, se consolida un marco holístico que combina elementos conceptuales, técnicos y prácticos para una comprensión y atención integral de la problemática, (Ministerio de Educación Nacional, s.f.).

El machine learning o aprendizaje automático como nos dice (Ramirez D, 2018) se refiere a un conjunto de técnicas y métodos que permiten que los sistemas informáticos adquieran conocimientos y capacidades a partir del procesamiento de grandes volúmenes de datos. En lugar de ser programados explícitamente con un conjunto de instrucciones predefinidas, estos sistemas tienen la habilidad de detectar patrones complejos, extraer insights relevantes y generar modelos predictivos de manera autónoma mediante algoritmos de entrenamiento especializados.

La principal utilidad del aprendizaje automático radica en su capacidad para abordar desafíos que involucran la exploración y análisis de cantidades masivas de información, descubriendo relaciones subyacentes que serían prácticamente imposibles de identificar y procesar manualmente. Estos enfoques resultan especialmente valiosos en áreas que requieren toma de decisiones basadas en datos, pronósticos precisos, reconocimiento de patrones, detección de anomalías, agrupamiento de observaciones afines y optimización de procesos complejos.

El machine learning encuentra aplicaciones en una amplia gama de dominios, desde el procesamiento de lenguaje natural y la visión por computadora, hasta sistemas de recomendación. La clave reside en aprovechar la potencia computacional y los avanzados modelos matemáticos para extraer conocimientos accionables a partir de los datos, superando las limitaciones de los enfoques tradicionales basados en reglas estáticas predefinidas.

Estas capacidades analíticas permiten a las organizaciones tomar decisiones más informadas, optimizar recursos, mejorar la eficiencia operativa y brindar experiencias más personalizadas a sus usuarios o clientes. En resumen, el aprendizaje automático representa una herramienta poderosa para desbloquear el valor oculto en los datos y automatizar tareas complejas con un alto grado de precisión y escalabilidad.

Para la estrategia de clasificación usaríamos algoritmos de aprendizaje supervisado donde se podrían usar etiquetas como por ejemplo (retirado-no retirado) en cada pila de datos (estudiantes) dándonos variables para predecir estos retiros o en su defecto dándonos características varias para luego tomar acciones al respecto. Teniendo también en cuenta todos los datos dados por (Ministerio de Educación Nacional, s.f.).

Teniendo uno muy conocido el cual sería el (SMV) máquinas de vectores de soporte que involucran directamente los siguientes datos:

1. Recopilar un conjunto de datos históricos con información de estudiantes, incluyendo su estado final (retirado-no retirado) y variables relevantes como rendimiento académico, datos socioeconómicos, participación en programas, etc.
2. Dividir aleatoriamente estos datos en conjuntos de entrenamiento y prueba.

Escalar o normalizar las variables predictoras para evitar que aquellas con mayor magnitud dominen el modelo.

3. Entrenar el algoritmo SVM con el conjunto de entrenamiento, buscando encontrar el hiperplano óptimo que clasifique mejor las observaciones.
4. Ajustar los parámetros del modelo (como el parámetro de penalización y el kernel) mediante técnicas como validación cruzada.
5. Evaluar el desempeño del modelo SVM en el conjunto de prueba

6. Interpretar los términos asignados a cada variable para comprender su importancia relativa en el fenómeno de deserción.
7. Utilizar el modelo entrenado para clasificar nuevas instancias y predecir la probabilidad de que un estudiante abandone sus estudios.

Otras alternativas de clasificación podrían ser árboles de decisión, regresión logística, naive bayes, redes neuronales, entre otros. La elección dependerá de factores como el tamaño y calidad de los datos disponibles, requisitos de interpretabilidad, tasas deseadas de precisión y otros criterios de rendimiento.

Pregunta problema:

¿Cómo desarrollar una estrategia computacional para ayudar en la identificación temprana de casos de deserción escolar, a partir de algoritmos de Machine Learning?

La deserción escolar es un problema complejo que afecta a los sistemas educativos de todo el mundo. La identificación temprana de estudiantes en riesgo de deserción es crucial para implementar medidas de intervención y apoyo oportunas. El Machine Learning (ML) ofrece un conjunto de herramientas prometedoras para desarrollar estrategias computacionales que automaticen la identificación temprana de casos de deserción.

Metodología:**1. Recopilación de datos:**

Datos históricos de estudiantes: Matrícula, asistencia, calificaciones, comportamiento, participación en actividades extracurriculares, etc.

Datos socioeconómicos: Ingreso familiar, nivel educativo de los padres, acceso a internet, etc.

Datos contextuales: Ubicación de la escuela, zona a donde pertenece, etc.

2. Preprocesamiento de datos:

Limpieza de datos: Eliminar valores nulos, inconsistentes o duplicados.

Normalización de datos: Escalar las variables a una escala común.

Imputación de datos: Llenar los valores faltantes con métodos apropiados.

3. Selección de algoritmos de ML:

Algoritmos de clasificación: Regresión logística, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Random Forest, etc.

Algoritmos de detección de anomalías: Local Outlier Factor (LOF), Isolation Forest, etc.

4. Entrenamiento y evaluación de modelos:

Dividir los datos en conjuntos de entrenamiento y prueba.

Entrenar los modelos de ML con el conjunto de entrenamiento.

Evaluar el rendimiento de los modelos con el conjunto de prueba.

Optimizar los parámetros de los modelos para obtener la mejor precisión posible.

5. Implementación y monitoreo:

Implementar el modelo seleccionado en un sistema computacional.

Monitorear el rendimiento del modelo en producción y realizar ajustes cuando sea necesario

Consideraciones éticas:

Asegurar la privacidad y confidencialidad de los datos de los estudiantes.

Implementar medidas para evitar el sesgo algorítmico y la discriminación.

Brindar a los estudiantes y a sus familias información clara y transparente sobre el uso de ML para la identificación temprana de la deserción escolar.

Acercamiento a los datos:

Los datos que usaremos a continuación son tomados de (Ministerio de Educación Nacional, s.f.). Ya que nos permiten tener una visión mas amplia de sus estudiantes de preescolar, básica primaria, secundaria y universitaria. A través del sistema integrado de matricula SIMAT ya que esta plataforma tiene registro de matrícula, asistencia y por supuesto el rendimiento académico de cada estudiante

Descripción de variables.

Acá tenemos muchas variables las cuales hemos reducido de manera clara para poder dar un detalle desde 2 puntos de vista, uno general (nivel país) y uno más específico (nivel Departamentos)

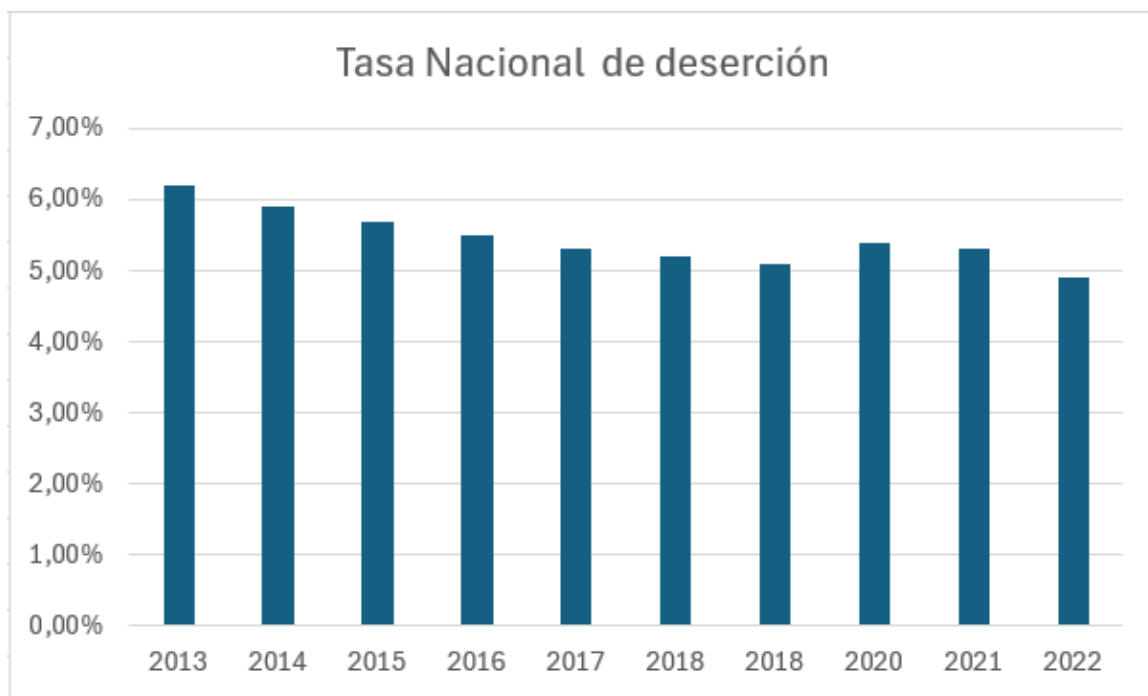
Variables tales como genero (niño o niña), año, zona urbana, zona rural, departamentos, primaria, media, secundaria, tasa de deserción, tasa de deserción de niñas, tasa de deserción de niños.

Posibles aplicaciones.

Estos datos nos pueden proporcionar mucha ayuda para tomar decisiones a un futuro o para reforzar las medidas que ya tenemos implementadas, ya que tal vez necesiten un reajuste, como por ejemplo implementar becas y ayudas económicas para estudiantes de bajo recursos de una manera mas eficiente, fortalecer la educación rural para mejorar su calidad educativa también en términos de infraestructura podríamos identificar donde hay falencias, mejorar el apoyo psicológico a los estudiantes y también tener estrategias para revertir el trabajo infantil que es muy común y el embarazo a temprana edad.

Aproximaciones con gráficos.

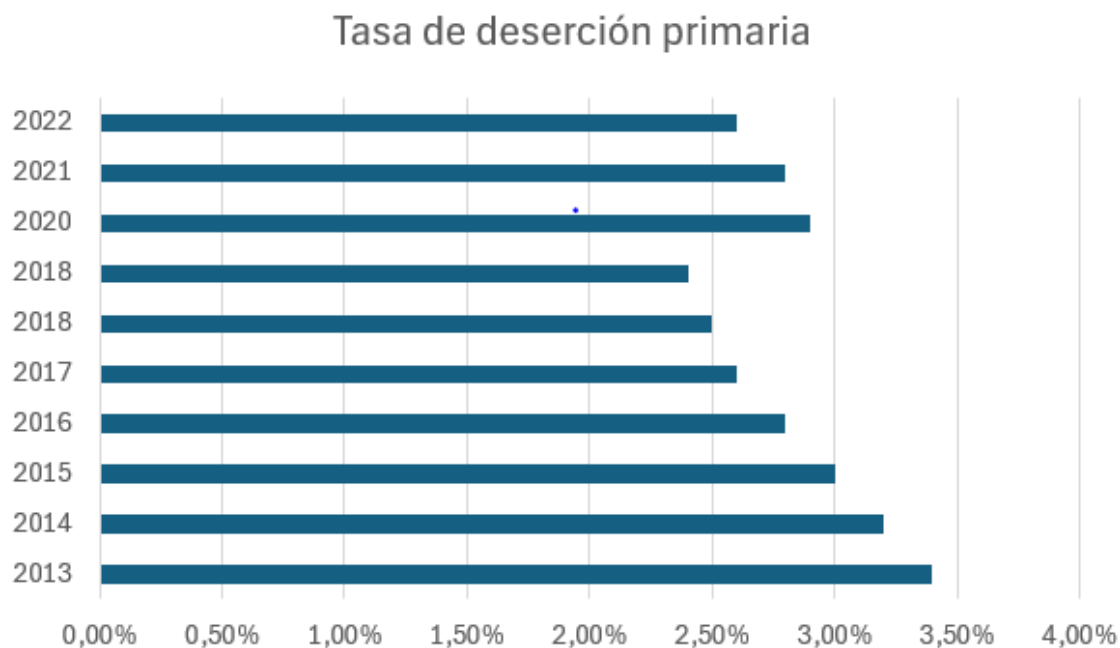
Grafica 1



Autor

En esta imagen podemos ver la tasa de deserción de los últimos 10 años, claramente podemos ver que ha disminuido, tengamos en cuenta que esto es a nivel general del país y acá tenemos es una gráfica directamente de deserción (no se toma en cuenta si tal vez el alumno gana el año o tal vez si lo llega a perder)

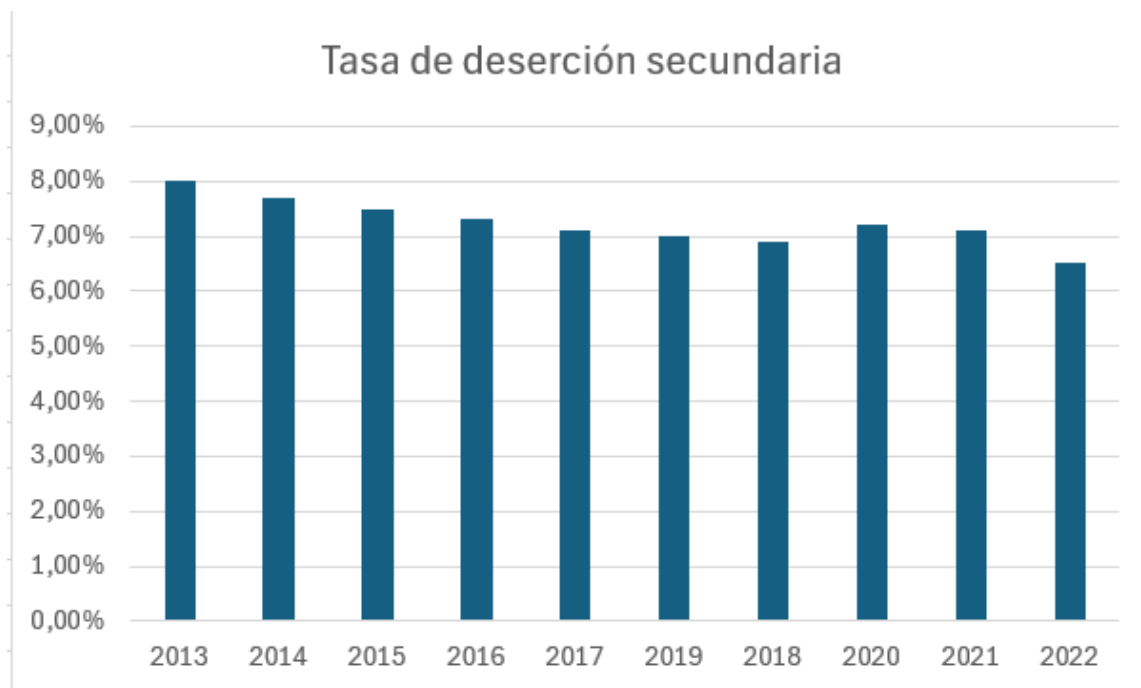
Grafica 2



Autor

La anterior es la tasa de deserción de los estudiantes de primaria donde en el 2020 se noto un leve incremento ya que recordemos la situación padecida a nivel mundial (COVID-19), se dio ya que muchos padres de familia tenían miedo a que sus hijos no aprendieran lo suficiente o en su defecto fuesen infectados con el virus ya que al finalizar el año no los reincorporaron.

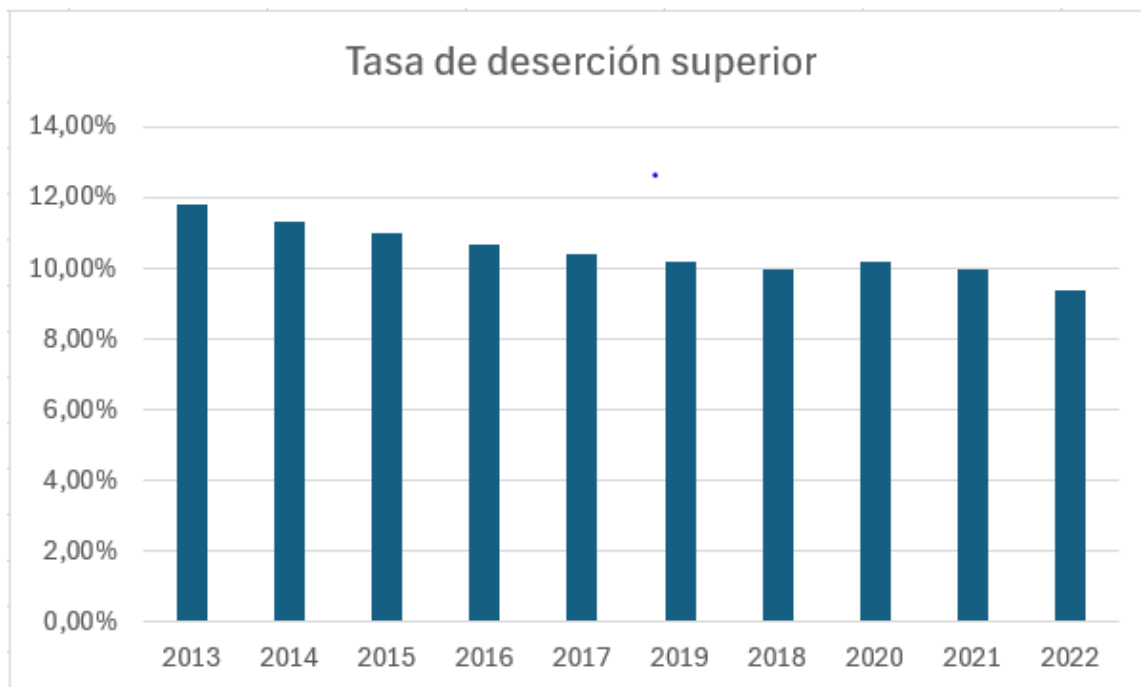
Grafica3



Autor

Acá tenemos la grafica de secundaria, vemos que se mantiene entre un 6% y un 8% en los últimos 10 años no hay gran cambio, esto se da ya que muchos son obligados a trabajar según los datos proporcionados por (Ministerio de Educación Nacional, s.f.).

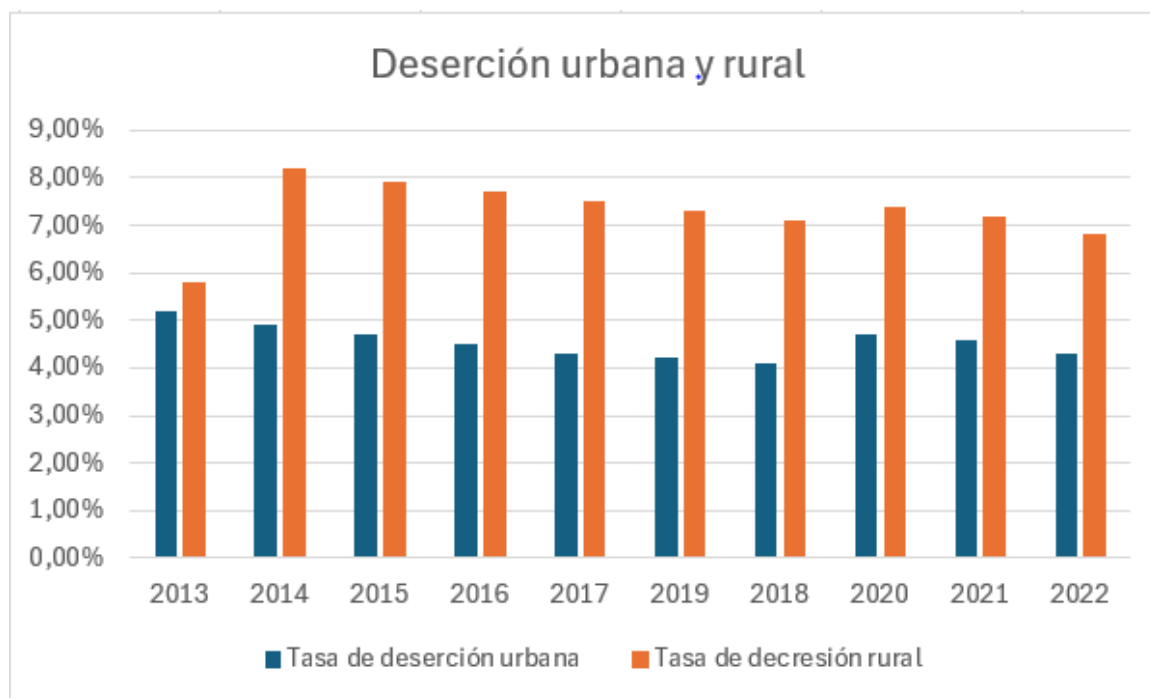
Grafica 4



Autor

Tenemos la tasa de deserción superior, la cual también se da por varios factores como pueden ser, altos costos en las carreras, poco campo o tal vez el nulo campo laboral y las oportunidades que presenta el país para desarrollarse como profesional.

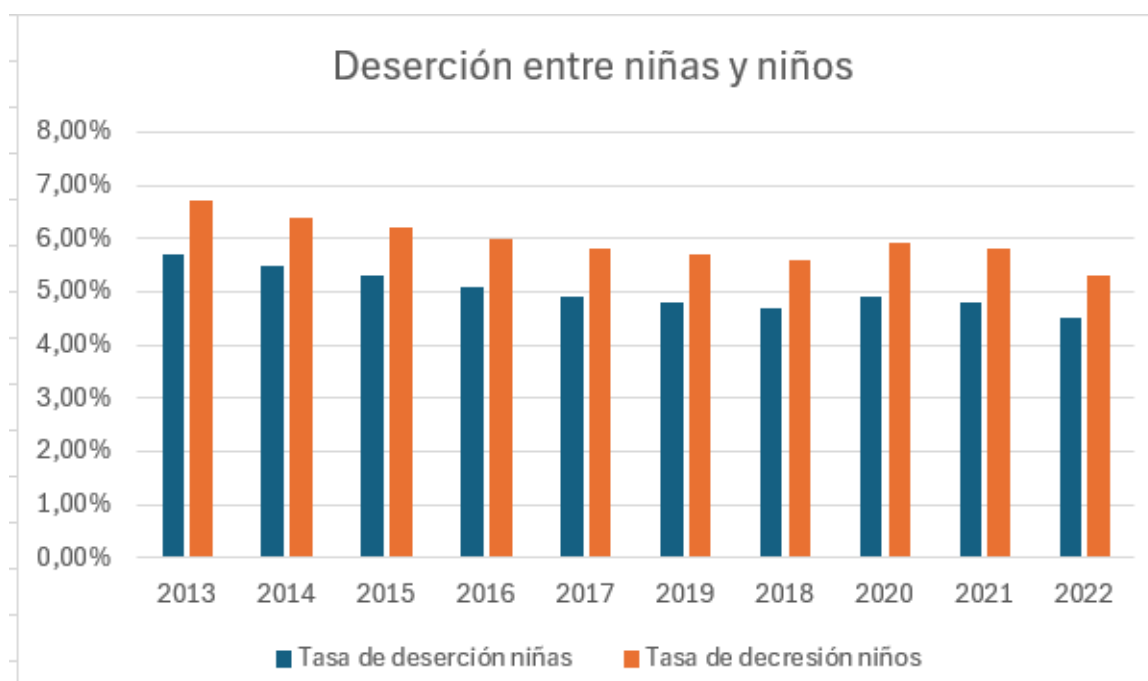
Grafica 5



Autor

Acá tenemos una comparación de deserción de zona urbana con zona rural vemos que las zonas rurales son mucho mayores a las urbanas, se debe a muchos factores tales como seguridad, compromiso por parte de padres de familia, también infraestructura de las instituciones, material de estudio y compromiso por parte nacional.

Grafica 6



Autor

Acá tenemos la comparativa entre niñas y niños donde nos damos cuenta que son mas los niños que se retiran que las niñas (tomando en cuenta esto hasta secundaria). Ya los factores se han mencionado anteriormente en las demás graficas.

Todo lo anterior fue hecho con ayuda de código en Python elaborado con la base de datos ya estructurada y organizada tomada del (Ministerio de Educación Nacional, s.f.).

Objetivos:**Objetivo general.**

Implementar un algoritmo computacional para el análisis y toma de decisiones a partir de datos de Deserción escolar, utilizando estrategias de machine learning.

Objetivos específicos.

- Caracterizar y procesar los datos de interés, con miras a la toma de decisiones informadas.
- Implementar un algoritmo de Machine learning para la toma de decisiones a partir de los datos de interés.
- Evaluar y analizar el desempeño de los algoritmos implementados para la toma de decisiones.
- Validar el funcionamiento de toma de decisiones a partir de datos nuevos.

Desarrollo e implementación del aprendizaje

Primero organizamos los datos, ya que estamos creando un método efectivo para evitar deserciones, este nos ayudaría a tener todo bajo cierto control y de una manera organizada, ya que es un método supervisado, con él, podremos mirar patrones y establecer unas reglas basadas en lo que se vaya demostrando para darnos cuenta de que

es lo que provocaría la deserción y así poder evitarla, en cada análisis nos dimos cuenta la cantidad en porcentaje (%), de estudiantes que se retiraban sin culminar el año escolar, ayudándonos mediante la separación de estos lograr algo mas personalizado y mas detallado con lo trabajar.

Procesamiento de los datos

Al tomar los datos nos damos cuenta de que estaban todos los datos mezclados, si son todos los datos de deserciones, pero debemos saber que no todos tienen las mismas necesidades.

Primero que todo toco separar varias cosas como vimos anteriormente en las gráficas, eliminé varias filas y columnas ya que no aportaban un impacto directo en las gráficas ni a lo que personalmente se quería llegar, tales como:

Código del departamento ya que nos centramos en Antioquia y en general del país, tamaño del territorio, población en dicho territorio, la tasa de matriculación, también se realizó un consenso de manera general por años, para no trabajar por meses lo hicimos al año, ya que mucho que se desvinculaban de una institución podrían vincularse a otra al poco tiempo según los datos suministrados, solo queremos saber quienes continuaron la actividad académica, quitamos también las sedes conectadas a internet, por eso mismo hicimos la división de zona urbana y zona rural, también omitimos los que ganaron el año académico y también los estudiantes que perdieron y todo lo que esta conlleva como promociones al grado siguiente y repitencia. Al tener todo esto separado ya podemos trabajar con la problemática de manera directa sin tener datos flotantes y sin tener datos repetidos ya que había filas y columnas duplicadas.

Modelo de toma de decisiones

Se realizo de manera clasificatoria, porque tenemos que saber de que manera vamos a actuar y con qué tipo de población lo vamos a hacer, bien sea con básica primaria, básica secundaria o superior.

Al tener las filas y columnas que se necesitaban, debemos tener en cuenta que la problemática de los estudiantes universitarios no son las mismas que la de los estudiantes de básica primaria, lo que se busca con esto es tener una mayor claridad de las situaciones de dicha población para comenzar a dar soluciones como las que se han planteado a lo largo de este documento ya que podemos ver con las claridad y comenzar incluso a ser mas clasificaciones por ejemplo, si es que hay mujeres adolescentes embarazadas o si tal vez los estudiantes están sufriendo bullying de algún tipo ya sea por su religión u orientación sexual y demás, poderlas plantear, separar, tener clara la problemática y luego comenzar a darle solución.

Implementación en contextos reales


Con los datos nuevos se realiza de la siguiente manera. Se ponen los datos directamente

```

import pandas as pd
from google.colab import files
uploaded = files.upload()
for filename in uploaded.keys():
    Dataset = pd.read_excel(filename)

```

Dataset.head(11)

 Elegir archivos: Sin archivos seleccionados Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving final.xlsx to final (5).xlsx

	Año	Tasa Nacional de Deserción	Tasa Deserción Primaria	Tasa Deserción Secundaria	Tasa Deserción Media	Tasa Deserción Urbana	Tasa Deserción Rural	Tasa Deserción Niñas	Tasa Deserción Niños
0	2013	6.2%	3.4%	8.0%	11.8%	5.2%	8.5%	5.7%	6.7%
1	2014	5.9%	3.2%	7.7%	11.3%	4.9%	8.2%	5.5%	6.4%
2	2015	5.7%	3.0%	7.5%	11.0%	4.7%	7.9%	5.3%	6.2%
3	2016	5.5%	2.8%	7.3%	10.7%	4.5%	7.7%	5.1%	6.0%
4	2017	5.3%	2.6%	7.1%	10.4%	4.3%	7.5%	4.9%	5.8%
5	2018	5.2%	2.5%	7.0%	10.2%	4.2%	7.3%	4.8%	5.7%
6	2019	5.1%	2.4%	6.9%	10.0%	4.1%	7.1%	4.7%	5.6%
7	2020	5.4%	2.9%	7.2%	10.2%	4.7%	7.4%	4.9%	5.9%
8	2021	5.3%	2.8%	7.1%	10.0%	4.6%	7.2%	4.8%	5.8%
9	2022	4.9%	2.6%	6.5%	9.4%	4.3%	6.8%	4.5%	5.3%

Autor

De acá podemos usar cualquiera de los comandos que deseemos para poder realizar la implementación del trabajo y también nos aseguramos de comprobar los datos.



#Información de la estructura de datos

Conjunto_Datos.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 9 columns):
#   Column                                     Non-Null Count  Dtype
---  ---
0   Año                                         10 non-null     int64
1   Tasa Nacional de Deserción                 10 non-null     object
2   Tasa Deserción Primaria                   10 non-null     object
3   Tasa Deserción Secundaria                 10 non-null     object
4   Tasa Deserción Media                       10 non-null     object
5   Tasa Deserción Urbana                     10 non-null     object
6   Tasa Deserción Rural                      10 non-null     object
7   Tasa Deserción Niñas                      10 non-null     object
8   Tasa Deserción Niños                      10 non-null     object
dtypes: int64(1), object(8)
memory usage: 848.0+ bytes
```

Autor

Como ya teníamos todo de manera simplificada y lo que necesitábamos un pequeño ejemplo para quitar nombres es de esta manera

```
#Quitando columnas indeseadas
Conjunto_Datos=Conjunto_Datos.drop(['Tasa Deserción Media'],axis=1)
#resumen de los datos
Conjunto_Datos.head()
```

Año	Tasa Nacional de Deserción	Tasa Deserción Primaria	Tasa Deserción Secundaria	Tasa Deserción Urbana	Tasa Deserción Rural	Tasa Deserción Niñas	Tasa Deserción Niños
0 2013	6.2%	3.4%	8.0%	5.2%	8.5%	5.7%	6.7%
1 2014	5.9%	3.2%	7.7%	4.9%	8.2%	5.5%	6.4%
2 2015	5.7%	3.0%	7.5%	4.7%	7.9%	5.3%	6.2%
3 2016	5.5%	2.8%	7.3%	4.5%	7.7%	5.1%	6.0%
4 2017	5.3%	2.6%	7.1%	4.3%	7.5%	4.9%	5.8%

Autor

Acá se quito para dar un ejemplo el “Tasa Deserción Media”

```

▶ # Identificar y eliminar duplicados
print('Número de filas antes de eliminar duplicados:', len(Conjunto_Datos))
Conjunto_Datos = Conjunto_Datos.drop_duplicates()
print('Número de filas después de eliminar duplicados:', len(Conjunto_Datos))

```

```

Número de filas antes de eliminar duplicados: 10
Número de filas después de eliminar duplicados: 10

```

Autor

Se eliminan las líneas duplicadas o los elementos que estén duplicados

Para analizar los gráficos anteriores.

Y así podríamos avanzar para poder obtener todos los datos que necesitemos claro esta con comando necesarios, por ejemplo para poder tener el algoritmo de clasificación seria.

```

▶ # Importar librerías necesarias
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, classification_report
|

```

Autor

También tenemos que tener en cuenta como tenemos guardado el archivo

```

▶ # Cargar datos |
datos = pd.read_csv('datos_final.csv')

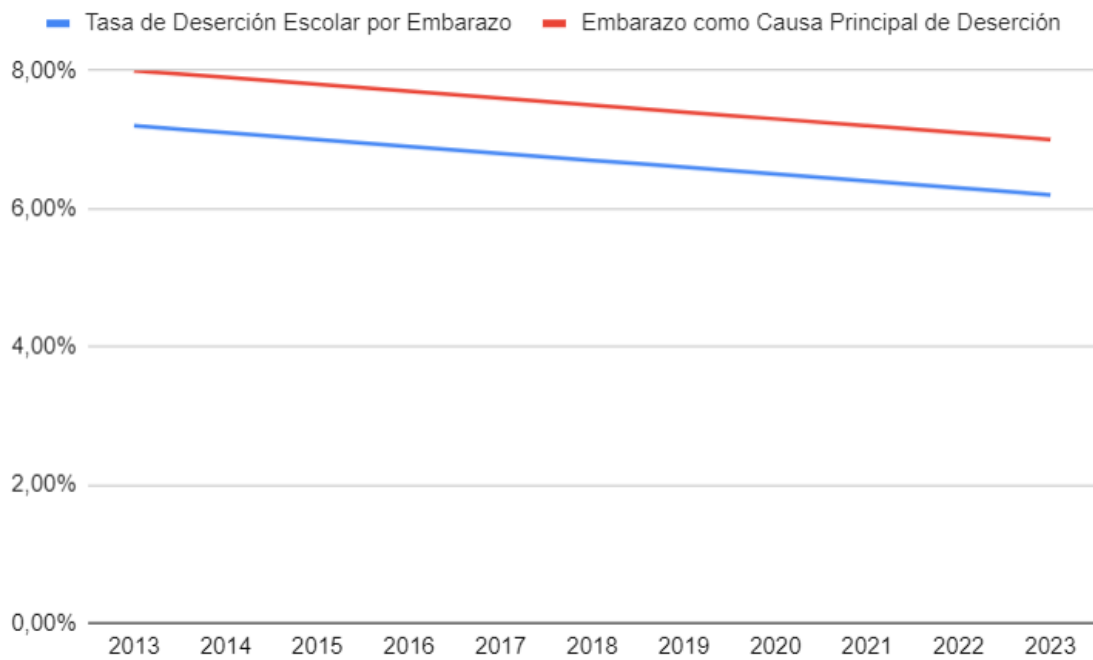
```

Autor

Resultados adicionales

Las únicas modificaciones realizadas fueron, eliminar las filas y columnas ya mencionadas, pero agregando un par donde nos dicen deserción por mujeres embarazadas.

Figuras y tablas



Autor

Se presenta una tabla de deserción por motivos de embarazo en las mujeres donde se nota una gran disminución a lo largo de 10 años

Año	Tasa Nacional de Deserción	Tasa Deserción Primaria	Tasa Deserción Secundaria	Tasa Deserción Media	Tasa Deserción Urbana	Tasa Deserción Rural	Tasa Deserción Niñas	Tasa Deserción Niños
2013,00	6.2%	3.4%	8.0%	11.8%	5.2%	8.5%	5.7%	6.7%
2014,00	5.9%	3.2%	7.7%	11.3%	4.9%	8.2%	5.5%	6.4%
2015,00	5.7%	3.0%	7.5%	11.0%	4.7%	7.9%	5.3%	6.2%
2016,00	5.5%	2.8%	7.3%	10.7%	4.5%	7.7%	5.1%	6.0%
2017,00	5.3%	2.6%	7.1%	10.4%	4.3%	7.5%	4.9%	5.8%
2018,00	5.2%	2.5%	7.0%	10.2%	4.2%	7.3%	4.8%	5.7%
2019,00	5.1%	2.4%	6.9%	10.0%	4.1%	7.1%	4.7%	5.6%
2020,00	5.4%	2.9%	7.2%	10.2%	4.7%	7.4%	4.9%	5.9%
2021,00	5.3%	2.8%	7.1%	10.0%	4.6%	7.2%	4.8%	5.8%
2022,00	4.9%	2.6%	6.5%	9.4%	4.3%	6.8%	4.5%	5.3%

Fuente: (Ministerio de Educación Nacional, s.f.).

Base de datos utilizada (recordemos que esta ya esta reducida y se enfoca en lo que necesitamos)

Conclusiones

Después de hacer los gráficos y analizar la tabla obtenida mediante la Secretaría Educación Nacional, nos damos cuenta que en términos generales se ha reducido la deserción en 6.2% hasta un 4.2%.

Tenemos brechas en nuestro nivel educativo entre mayor es el nivel más deserción nos encontramos.

Diferencias de género, aunque estas son leves hay más deserción en niños que en niñas, ya luego en un sistema superior las niñas están casi a la par por el tema de los embarazos

Referencias

Zhou, Z. H. (2021). Machine learning. Springer nature.

Hinestroza Ramírez, D. (2018). El Machine Learning a través de los tiempos, y los aportes a la humanidad.

Ministerio de Educación Nacional de Colombia. (2014). Indicadores educativos: Enero 2014. [Archivo PDF]. Bogotá

Valencia, J. H. (2017). Análisis de la deserción estudiantil en la FCECEP utilizando Machine Learning específicamente mapas auto organizados de Kohonen. Santiago de Cali

Colombia Potencia De La Vida. (15 de 03 de 2022). Colombia Potencia De La Vida. Obtenido de Colombia Potencia De La Vida: <https://www.mineducacion.gov.co/portal/Preescolar-basica-y-media/Sistema-de-educacion-basica-y-media/233839:Sistema-educativo-colombiano>

Shinde, P. P., & Shah, S. (2018, August). A review of machine learning and deep learning applications. In *2018 Fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-6). IEEE.