**Planteamiento del proyecto**

* Idea del proyecto:

El objetivo de este proyecto es realizar ingeniería inversa sobre un modelo de aprendizaje automático. Este modelo se centra en medir la probabilidad de satisfacción con respecto a las calificaciones obtenidas en el examen de estado nacional de Colombia, específicamente en la ciudad de Bogotá. Utilizaremos datos demográficos recopilados por el ICFES de los participantes que realizaron el examen en el año 2019.

* Objetivo principal:

Nuestro principal objetivo es comprender exhaustivamente el proceso de trabajo y el funcionamiento del proyecto. Esto se logrará mediante la utilización de diagramas UML y un análisis detallado, permitiéndonos abordar el proyecto de manera más profesional y técnica.

* Objetivos Específicos:

Realizar un análisis detallado de la relación entre los datos demográficos y los resultados del examen del ICFES. Este análisis comprenderá una evaluación exhaustiva de cada etapa del proyecto, con el propósito de identificar los componentes más significativos y comprender sus relaciones y estructura, así como las interacciones entre ellos. Además, se presentarán los resultados obtenidos y se propondrán mejoras o formas de optimización, en caso de ser necesario.

* Planteamiento del problema:

En Colombia, se administra un examen de estado conocido como ICFES a los bachilleres, abarcando diversas áreas de conocimiento como matemáticas, ciencias sociales, lectura crítica, ciencias naturales e inglés, según los lineamientos establecidos por el gobierno colombiano. Este examen se aplica a todos los bachilleres del país, quienes inicialmente responden preguntas demográficas que proporcionan información sobre su estilo de vida, hábitos alimenticios, origen y otros datos relevantes. Esta información se considera crucial para investigaciones que buscan establecer una relación entre el contexto socioeconómico del estudiante y su satisfacción académica. El objetivo es mejorar la calidad de vida del estudiante para favorecer su desarrollo y aprendizaje óptimos.

* Necesidades:

Este proyecto específico demanda un profundo conocimiento en metodologías de trabajo para la ciencia de datos, como el CRISP-DM, así como habilidades en la creación de diagramas en lenguajes como Python. Además, se requiere un entendimiento sólido de las herramientas y librerías utilizadas en el desarrollo del proyecto.

* Estado del arte:

Se aprovechará otro proyecto elaborado por estudiantes de análisis bibliográfico, donde se recopilará documentación sobre las herramientas utilizadas en este proyecto y estudios previos relacionados con el tema. El objetivo es obtener un entendimiento exhaustivo que nos permita abordar de manera óptima el proyecto actual. Además, se llevará a cabo un análisis detallado de metodologías como CRISP-DM, SEMMA, ASUM DM y TDPS, con el fin de identificar cuál se adapta mejor a las necesidades y objetivos del proyecto en curso.

* Alcance del proyecto:

El objetivo es explorar una nueva área de estudio: el impacto de la alimentación en el aprendizaje de los estudiantes. Buscamos comprender en profundidad este tema para poder proponer cambios en el sistema educativo actual, especialmente en las escuelas que deberían priorizar el bienestar de los niños en todas sus dimensiones, incluyendo su alimentación. Nuestra meta es promover las mejores prácticas en estos entornos, asegurando así una contribución significativa al desarrollo y conocimiento futuro de los estudiantes.

**Requisitos**

* Recolección de Requisitos:

El proyecto se basa en la observación y curiosidad respecto a las variables sociodemográficas en Colombia y su relación con la satisfacción académica en el examen de estado ICFES. Utilizando datos de acceso público, la investigación se enfoca en entender cómo estas variables influyen en el rendimiento y la satisfacción de los estudiantes. Esta curiosidad está motivada por el sistema de becas en Colombia, estrechamente vinculado a los resultados del ICFES. Las preguntas clave que guían esta investigación son:

1. ¿Existen estratos sociales que tienen mayores oportunidades de acceder a estos beneficios?
2. ¿Existen estratos sociales que tienen mayores oportunidades de acceder a estos beneficios?
3. ¿Existen estratos sociales que tienen mayores oportunidades de acceder a estos beneficios?

A través de la exploración de datos, se espera formular nuevas preguntas que puedan ser respondidas con los mismos datos obtenidos, o en su defecto, sentar las bases para futuras investigaciones.

* Historias de usuario

* Documentos adicionales para el análisis

Tal como se ha mencionado en diversas partes del documento, se utilizará la base de datos pública del ICFES del año 2019. Para el análisis y limpieza de estos datos, se empleará el lenguaje de programación Python, aprovechando sus librerías especializadas en manipulación y análisis de datos.