**Proyecto**

**Nombre:** Juan sebastian perdomo Becerra [js.perdomo@unal.edu.co](mailto:js.perdomo@unal.edu.co)

Detección Temprana: Proyecto de Alerta en Enfermedades Mentales con modelos predictivos

1. **Introducción**

En un mundo donde la salud mental se ha convertido en una prioridad esencial, la capacidad de detectar enfermedades mentales de manera temprana es crucial para proporcionar la atención adecuada a quienes la necesitan. El proyecto "Detección Temprana" busca implementar un sistema innovador que, mediante el análisis de fuentes de datos relacionadas con la salud mental, permita identificar patrones y factores de riesgo en la población. Esto no solo contribuirá a la comprensión de estas enfermedades, sino que también facilitará la intervención oportuna, mejorando así los resultados de salud.

A lo largo de las últimas décadas, la prevalencia de enfermedades mentales ha aumentado de manera alarmante en todo el mundo. Según la OMS, una de cada cuatro personas se verá afectada por un trastorno mental a lo largo de su vida. En 2019, se estimó que más de 970 millones de personas padecían condiciones mentales como depresión, ansiedad y trastornos bipolares, siendo esta última pandemia amplificada por la crisis sanitaria global provocada por la COVID-19. Además, se estima que solo el 36% de las personas que tienen trastornos mentales reciben atención adecuada.

1. **Contexto del Problema:**

La creciente prevalencia de enfermedades mentales a nivel mundial, sumada a la falta de detección temprana y acceso a tratamientos adecuados, representa una crisis de salud pública. Los sistemas de salud actuales enfrentan desafíos significativos para identificar a las personas en riesgo y brindarles la atención necesaria de manera oportuna. La ausencia de herramientas predictivas basadas en datos limita la capacidad de los profesionales de la salud para anticiparse a la aparición de trastornos mentales y diseñar intervenciones personalizadas. Este proyecto busca abordar esta problemática al desarrollar un sistema de alerta temprana capaz de identificar patrones y factores de riesgo asociados con enfermedades mentales, con el objetivo de mejorar la detección y el manejo de estas condiciones.

1. **Alcance del Proyecto:**

El proyecto se centrará en el desarrollo e implementación de un modelo predictivo capaz de identificar individuos en riesgo de desarrollar enfermedades mentales a partir del análisis de diversas fuentes de datos relacionadas con la salud

* Análisis descriptivo de las enfermedades, características y estadística de las diferentes categorías que se pueden tener.
* Modelo predictivo, con algoritmos supervisados que a partir de un input se puede estimar la probabilidad de presentar una enfermedad

1. Descripción de los datos

El dataset se encuentra dispuesto en el Github que se utiliza para el proyecto, consta de 20 columnas y 93800 filas. El detalle de cada una de las columnas se puede ver en la siguiente tabla:

|  |  |  |
| --- | --- | --- |
| **Nombre columna** | **Dtype** | **Detalle** |
| id | int64 | Número de identificación |
| Name | object | Nombre de persona |
| Gender | object | Genero |
| Age | float64 | Edad |
| City | object | Ciudad de origen |
| Working Professional or Student | object | Estado profesional |
| Profession | object | Profesión de la persona |
| Academic Pressure | float64 | Escala de presión académica |
| Work Pressure | float64 | Escala de presión laboral |
| CGPA | float64 | Promedio escolar |
| Study Satisfaction | float64 | Escala de satisfacción |
| Job Satisfaction | float64 | Escala de satisfacción |
| Sleep Duration | object | Clasificación de horas de sueño (ej. 5-6 horas) |
| Dietary Habits | object | Hábitos alimenticios (saludable, moderado, no saludable |
| Degree | object | Titulo profesional |
| Have you ever had suicidal thoughts? | object | Pregunta de si o no |
| Work/Study Hours | float64 | Horas de trabajo |
| Financial Stress | float64 | Escala de presión o estrés financiero |
| Family History of Mental Illness | object | Una pregunta de si o no |
| Depression | int64 | Escala de depresión |

1. Métodos de recolección y almacenamiento

Estos datos fueron recolectados de diferentes hospitales y análisis que se hicieron a población de universidades. Estos datos en un futuro se pueden tener en un data lake y consultar con una API con la nueva información que vaya llegando

1. Herramientas y su uso

Para este ejercicio se tendrá la información en CSV con carga manual y se espera llevar a un servicio de Amazon Web service de S3

1. Estándares de metadatos

Toda la información recopilada en el marco de este proyecto será etiquetada con metadatos precisos y detallados, asegurando su trazabilidad y cumplimiento de las regulaciones en materia de protección de datos personales. Los datos serán almacenados en entornos de datos seguros y encriptados, cumpliendo con los más altos estándares de la industria y las normativas vigentes. Se establecerán mecanismos de acceso restringido y auditables para garantizar la confidencialidad de la información y evitar cualquier uso no autorizado. Los datos serán conservados durante un mínimo de siete años, en línea con las recomendaciones de las autoridades sanitarias y las mejores prácticas en el campo de la investigación en salud mental.

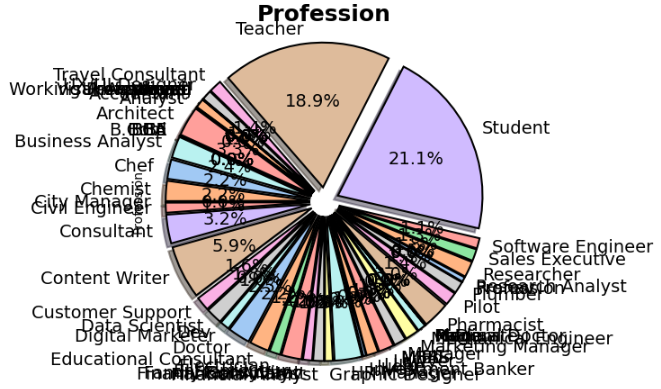
1. Estrategia de respaldo

Se utilizará una plataforma de almacenamiento en la nube altamente segura y escalable, como **Amazon S3 (Simple Storage Service)**, para almacenar los datos recopilados. Esta plataforma ofrece múltiples opciones de cifrado y autenticación, garantizando la confidencialidad de la información. Además, se implementarán copias de seguridad incrementales y diferenciales para minimizar el tiempo de recuperación en caso de desastre. El equipo técnico configurará y administrará la solución de almacenamiento en la nube, asegurando el cumplimiento de las mejores prácticas de seguridad y los requisitos regulatorios.

1. Exploración de los datos

Se limpia los resultados que están en blanco para tratar con información de pacientes que estén completa

Alguna de las profesiones que se presentan en los pacientes de enfermedades mentales



Comportamiento de las edades de los pacientes

Gráfico, Histograma

Descripción generada automáticamente

Profesión de las personas

Gráfico, Gráfico de barras

Descripción generada automáticamente

Segmentación entre estudiantes y trabajadores

Gráfico, Gráfico circular

Descripción generada automáticamente

Comportamiento del sueño

Gráfico, Gráfico de barras

Descripción generada automáticamente

Comportamiento de los comportamientos dietarios y distribución de géneros

Gráfico

Descripción generada automáticamente

Gráfico, Gráfico circular

Descripción generada automáticamente

1. Mock UP

Tablero de selección



Tablero de histórico



1. Reporte de trabajo

|  |  |
| --- | --- |
| **Nombre y Rol** | **Responsabilidad** |
| Diseñadora de la Interfaz de Usuario | - Diseño de la interfaz de usuario ycreación del Mock-up.- Previsualización del dashboard paraasegurar una experiencia de usuariointuitiva y funcional. |
| Ingeniero de Datos e Infraestructura | - Diseño de la infraestructura de datos.- Despliegue de los repositorios ygestión del versionamiento de losdatos.- Escalamiento de las aplicaciones queconformarán la APP para asegurar surendimiento y fiabilidad. |
| Analista de Datos y Modelado Predictivo | - Creación de los repositorios de DVC yGitHub.- Análisis descriptivo y predictivo de losdatos para entender patrones ytendencias. |