

Predicción del Riesgo de Hipertensión Mediante Minería de Datos.

Objetivos: Desarrollar un modelo predictivo para evaluar el riesgo de hipertensión utilizando técnicas de minería de datos.

- Implemente la metodología CRISP-DM para el desarrollo de este proyecto.
- Recolección de Datos.
 - Utilice los archivos CSV proporcionados: ensaantro2022_entrega_w.csv, Determinaciones_bioquímicas_cronicas_deficiencias_9feb23.csv, ensafisica2022_adultos_entrega_w.csv. Para mayores detalles de los dataset puede visitar la pagina <https://ensanut.insp.mx/encuestas/ensanutcontinua2022/descargas.php>
- Preprocesamiento de Datos.
 - Haga limpieza de los datos, manejo de valores faltantes, eliminación de duplicados.
 - Normalización/estandarización de los datos.
- Análisis Exploratorio de Datos (EDA).
 - Análisis estadístico para comprender las características de los datos.
 - Visualización de datos para identificar patrones y correlaciones.
- Correlación y Fusión de Datos.
 - Analice la correlación entre diferentes variables de los conjuntos de datos y elija las características más relevantes de cada dataset para hacer el análisis.
 - Fusione los datos en un solo conjunto de datos para el análisis.
- Agrupación con KMeans.
 - Implemente el algoritmo KMeans para identificar agrupaciones relevantes en los datos.
 - Analice los clústeres para identificar patrones relacionados con el riesgo de hipertensión.
- Desarrollo del Modelo Predictivo.
 - Utilice Gradiente Descendente Estocástico como algoritmo de optimización.
 - Implemente el modelo de regresión logística o lineal para hacer la clasificación, según considere sea el mas indicado.
- Evaluación del Modelo.
 - Evalúe el modelo utilizando métricas como precisión, sensibilidad, y el Área Bajo la Curva (AUC) del Receiver Operating Characteristic (ROC).
 - Interprete los resultados en el contexto del riesgo de hipertensión.
- Documentación y Comunicación de Resultados.
 - Redacte un artículo científico detallando el proceso y los hallazgos.
 - Incluya gráficos, tablas y análisis estadísticos en el artículo.
- Desarrollo de Código.
 - Desarrolle y documente el código en Python para cada paso del proyecto.
 - Asegúrese de que el código sea claro, bien comentado y reproducible.
- Conclusiones y Recomendaciones.
 - Resuma los hallazgos clave y su significado en la predicción del riesgo de hipertensión.
 - Proponga recomendaciones para futuras investigaciones o aplicaciones prácticas.
- **Sobre los productos finales.**
 - **Artículo Científico**, un documento detallado que presenta la investigación, metodología, análisis de datos, resultados, discusión y conclusiones, revisen el ejemplo (ee_formato_ieee.pdf).
 - Una carpeta en .zip que contenga todos los scripts utilizados en el análisis, debidamente documentados y organizados.