

Portafolio de análisis

Imanol Muñiz Ramírez A01701713

Link al repositorio con todas las evidencias de análisis:

https://github.com/ViejoAgrio/Portafolio_de_evidencias/tree/main/Analisis

Evidencias

SEG0403B - Módulo de reto

Verifica que los datos estén anonimizados para no violar normas o leyes de privacidad de información.

Especifica el proceso o estándar a seguir para validar el manejo de los datos y garantizar que solo el equipo tenga acceso a ellos en cumplimiento a las normativas

Investigación individual

- Link a evidencia:
https://github.com/ViejoAgrio/Portafolio_de_evidencias/blob/main/Analisis/Privacidad%20y%20Seguridad%20de%20los%20Datos.pdf
- Descripción de la evidencia: Se muestra la investigación individual sobre las leyes y normas que hay que tener en consideración para el proyecto y el plan para acatarlas con base en estándares internacionales.
- Correcciones: Únicamente se cambio a un formato no editable

Documento oficial del equipo

- Link a evidencia:
https://github.com/ViejoAgrio/Portafolio_de_evidencias/blob/main/Analisis/Pol%C3%ADtica%20de%20acceso.pdf
- Descripción de la evidencia: Se muestra el documento oficial del equipo sobre las leyes y normas que hay que tener en consideración para el proyecto y el plan para acatarlas con base en estándares internacionales.
- Correcciones: Se ajustó según la forma de trabajo y las investigaciones individuales.

Presenta registros claros sobre el seguimiento del proceso y las autorizaciones para tener los datos.

Bitácora de cambios

- Link a evidencia:
https://github.com/ViejoAgrio/Portafolio_de_evidencias/blob/main/Analisis/Bit%C3%A1cora%20de%20cambios.pdf
- Descripción de la evidencia: Se muestra la bitácora de cambios de nuestra base de datos.
- Correcciones: Solo se incrementaron los registros con el tiempo.

SMA0104C – Módulo Big Data y Estadística avanzada para la ciencia de datos

Trabaja con grandes volúmenes de datos.

Procesa grandes volúmenes de datos de manera eficiente.

Genera tableros útiles y correctos que apoyen a la toma de decisiones.

Workshop 5

- Link a evidencia:
https://github.com/ViejoAgrio/Portafolio_de_evidencias/blob/main/Analisis/Workshop5_ImanolMu%C3%B1izRamirez.ipynb
- Descripción de la evidencia: Este es el jupyter notebook del workshop 5 que consta de realizar modelos predictivos ARIMA – SARIMA de los productos internos brutos de las economías de México y Estados Unidos para finalmente hacer un modelo predictivo ARIMA – SARIMAX del PIB de México con variable exógena el cambio del PIB de Estados Unidos.
- Correcciones: No se hicieron correcciones.

Realiza los pasos adecuados para calibrar un modelo estadístico.

Mide correctamente el desempeño del modelo y sus métricas son correctas.

Interpreta los coeficientes del modelo, así como los resultados de las predicciones en el contexto del problema de manera correcta.

Utilización, procesamiento y visualización de grandes volúmenes de datos

- Link a evidencia:
https://github.com/ViejoAgrio/Portafolio_de_evidencias/blob/main/Analisis/NLP.ipynb
- Descripción de la evidencia: Se utiliza PySpark para procesar un dataset con un grande volumen de datos y generar un modelo de inteligencia artificial capaz de analizar textos para predecir el tema de su contenido.

- Correcciones:
 - Se justificaron más ampliamente las acciones llevadas a cabo
 - Se explicaron más ampliamente las observaciones del proceso
 - Se eliminaron prints innecesarios
 - Se verificó el código y se simplificó
 - Se añadieron más gráficos para una mejor comprensión de la situación y los resultados.
 - Se diversificaron los gráficos utilizados