

Trabajo final Inteligencia Artificial

Maestría IoT & IA

En este trabajo final los equipos deberán elegir 2 problemas a solucionar (no similares y utilizando métodos y algoritmos distintos para cada caso) aplicando algoritmos de Inteligencia Artificial. Algunos de los temas pueden ser los siguientes:

- Spam detection
- Product recommendation
- Fraud detection
- Document Classification
- Clasificación en general
- Series de tiempo para reconocimiento o predicción
- Instalación y entrenamiento de LLMs (Investigación por cuenta propia)
 - Instalar y entrenar LLM con documentos para asistencia de alumnos y profesores (reglamentos, normatividad y convocatorias, etc de la unison)
 - Para pacientes de diabetes
 - Para Pacientes de cáncer de mama
- Clasificación de frutas con modelo jerárquico (adaptación del trabajo hecho con el Prof. Benítez)
- Analisis de Sentimientos
- Filtro colaborativo

Bases de datos que pudieran ser de utilidad:

<https://www.kaggle.com/uciml/sms-spam-collection-dataset>

<https://www.kaggle.com/c/santander-product-recommendation/data>

<https://www.kaggle.com/mlg-ulb/creditcardfraud> (esta base de datos está desbalanceada, pues en la vida real los datos de fraude son menos del 1% de las transacciones. Deberán ser creativos para lidiar con esto).

<https://www.kaggle.com/sachinsharma1123/document-classification>

La solución es libre, pueden usar combinaciones de modelos (parte de la solución puede ser heurística si se justifica). Lo importante es modelar la realidad que les muestran los datos y dar una respuesta acorde que pueda escalar con el paso del tiempo y se sostenga.

Deberán entregar un informe y presentar sus resultados.

(04 diciembre): Presentar por equipo el planteamiento del problema, análisis de datos, gráficas y lo que han encontrado durante el análisis exploratorio.

Entrega de un reporte:

La estructura del reporte es libre, pero debe contener para cada problema:

- 1 Descripción del problema
- 2 Análisis exploratorio
- 3 Modelos - solución

- Pruebas y resultados
- Hallazgos más importantes
- Conclusiones

IMPORTANTE:

Estos problemas son COMPLEJOS. No esperen tener un buen desempeño en dos semanas (aunque si son diferenciables las clases lo tendrán). Lo que evaluaremos es la manera en que abordan el problema:

- Los criterios que toman para elegir la representación o features
- Los criterios que toman para elegir los algoritmos
- Lo exhaustivo del análisis exploratorio
- Los argumentos que justifiquen sus decisiones ante la toma de decisiones
- La calidad de la presentación, completitud y claridad del documento
- Aplica PCA? No en todos los casos es necesario
- Verificar si hay overfitting haciendo cross validation