

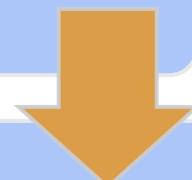
Sistema de Clasificación y Recomendación de Revistas Científicas

Ciencia de Datos en Ingeniería

Juan Antonio Hernández Rodríguez

Objetivos

Implementar un sistema capaz de asignar artículos a la revista correcta basándose en su contenido



Comparar el rendimiento entre métodos clásicos de aprendizaje automático y una aproximación conexionista (redes neuronales)



Establecer un motor de recomendación de artículos basado en el contenido semántico de los textos.

Datos utilizados

- **Origen:** Artículos de cuatro revistas de Elsevier: *Applied Ergonomics*, *Journal of Visual Communication*, *Neural Networks* y *Robotics and Autonomous Systems*.
- Composición del Dataset:
 - **Muestras:** Aproximadamente 6,000 artículos.
 - **Atributos:** Concatenación de **Título + Abstract + Palabras clave**
 - **Partición:** Estratificada 80% entrenamiento / 20% validación.



Modelos probados

LinearSVC, Regresión
Logistica y Naives Bayes
Multimodal

Red Neuronal BiGRU

Metodología

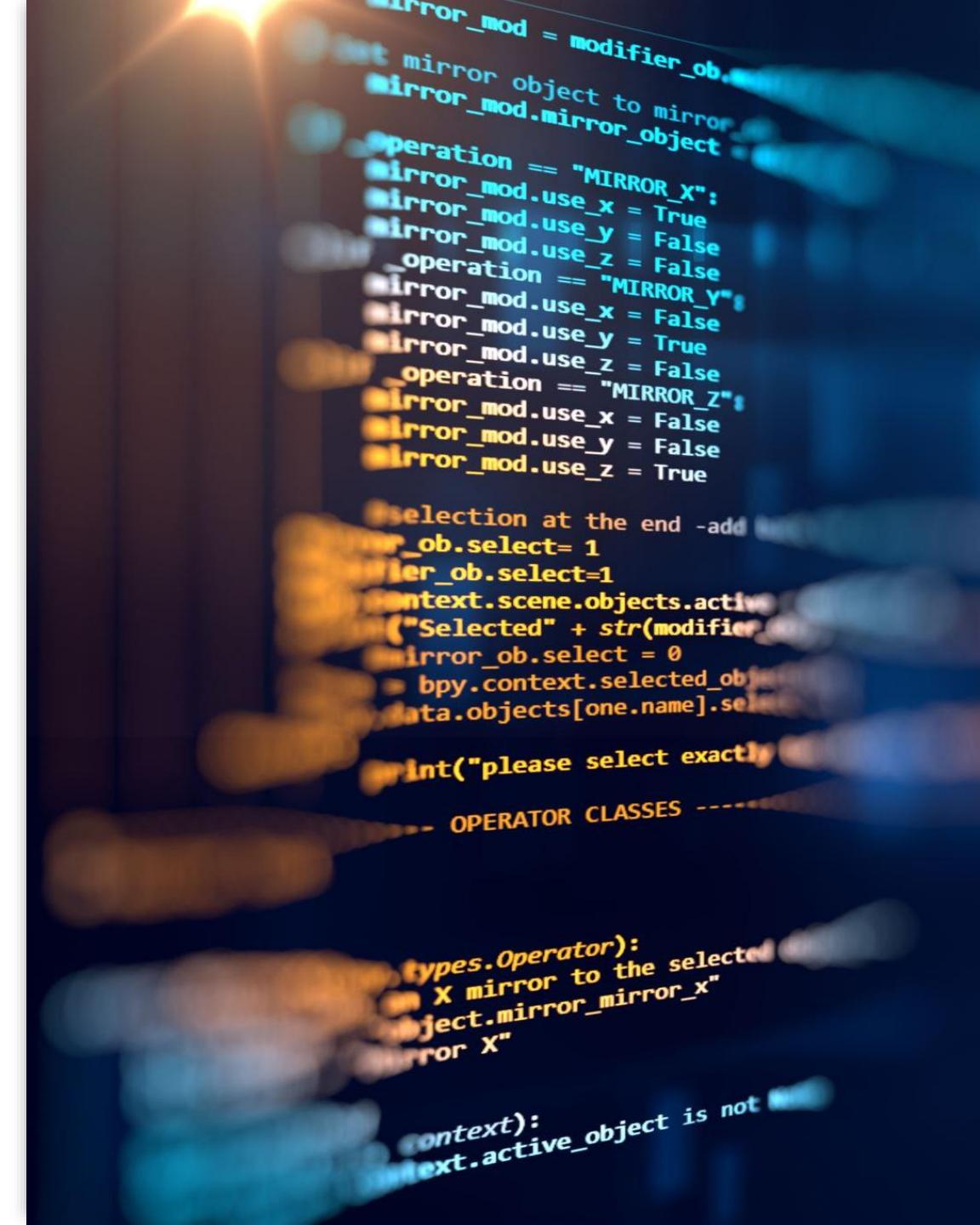
1. Construcción del Dataset: Recorrido y parseo de archivos JSON para extraer texto y etiquetas.

2. Extracción de Características:

- **Clásica:** Vectorización TF-IDF con parámetros optimizados.
- **Conexionista:** Tokenización y creación de vocabulario neuronal.

3. Entrenamiento:

- Uso de Sklearn para modelos lineales.
- Implementación en PyTorch con **AdamW** y **Early Stopping** para BiGRU



Resultados



El mejor modelo ha sido el LinearSVC con un 88.9% de Accuracy global.



BiGRU alcanzó un 83.5% de Accuracy, destacando por su capacidad de captura semántica profunda.



En cuanto a eficiencia, los modelos clásicos fueron drásticamente más rápidos (2.1s vs 184s de BiGRU).

Conclusiones

- Eficacia: Los métodos clásicos (LinearSVC) son superiores en precisión y eficiencia para este volumen de datos y clases.
- Potencial conexionista: BiGRU ofrece una alternativa robusta para capturar dependencias secuenciales complejas en abstracts largos.
- Trabajo Futuro: Explorar arquitecturas tipo Transformer o ampliar el sistema hacia recomendaciones personalizadas.