

# Investigación de algoritmos

## 1. Redes Neuronales Recurrentes (RNNs)

### Explicación del modelo:

- Las RNNs son diseñadas específicamente para procesar datos secuenciales o temporales. Su capacidad de "recordar" información previa dentro de una secuencia las hace útiles para tareas relacionadas con el lenguaje.
- En cada paso de tiempo, las RNNs utilizan el estado oculto previo junto con la entrada actual para producir una salida y actualizar su estado.

### Ventajas en la clasificación de relaciones biomédicas:

1. **Secuencialidad:** Capturan dependencias locales entre palabras consecutivas, lo que es importante para frases que describen relaciones entre entidades.
2. **Simplicidad:** Son fáciles de entrenar y adecuadas para proyectos iniciales con datos limitados.

### Limitaciones:

- **Dependencias largas:** Las RNNs tienen dificultades para capturar relaciones que dependen de palabras separadas por largas distancias.
- **Eficiencia:** No son paralelizables, lo que limita su escalabilidad en conjuntos de datos grandes.

### Justificación para uso:

- Son útiles si las relaciones entre entidades biomédicas dependen principalmente de palabras cercanas en los abstracts, pero su desempeño puede ser limitado si las relaciones son más complejas o distantes.

## 2. Long Short-Term Memory (LSTM)

### Explicación del modelo:

- Las LSTM son una mejora de las RNNs que introducen celdas de memoria y puertas de control para manejar mejor las dependencias a largo plazo en los datos secuenciales.

- Estas puertas deciden cuánta información debe olvidarse, almacenarse y usarse en cada paso de tiempo.

### **Ventajas en la clasificación de relaciones biomédicas:**

1. **Dependencias largas:** Las LSTM pueden capturar relaciones entre entidades que están separadas por muchas palabras o incluso frases.
2. **Contexto rico:** Son capaces de modelar secuencias más complejas y abstraer información relevante de largos abstracts biomédicos.
3. **Eficiencia:** Aunque no son tan rápidas como los Transformers, son más ligeras en términos de computación.

### **Limitaciones:**

- **Menor paralelismo:** No son tan rápidas como modelos basados en Transformers debido a su naturaleza secuencial.
- **Requieren ajustes:** Pueden necesitar una arquitectura cuidadosamente diseñada para manejar relaciones biomédicas.

### **Justificación para uso:**

- Son ideales si el proyecto involucra relaciones complejas entre entidades biomédicas que dependen tanto del contexto cercano como distante en los abstracts. Las LSTM son

## **3. Transformers (BERT como ejemplo)**

### **Explicación del modelo:**

- Los Transformers, como BERT, utilizan un mecanismo de **Self-Attention** que evalúa la importancia relativa de todas las palabras en la secuencia entre sí. Esto les permite capturar dependencias globales independientemente de la distancia entre palabras.
- BERT (Bidirectional Encoder Representations from Transformers) es un modelo preentrenado que entiende el contexto en ambas direcciones (izquierda y derecha).

### **Ventajas en la clasificación de relaciones biomédicas:**

1. **Relaciones globales:** Capturan relaciones complejas entre entidades, incluso si están separadas por largas distancias en los abstracts.
2. **Preentrenamiento:** BERT y sus variantes (SciBERT, BioBERT) pueden ser preentrenadas en grandes conjuntos de datos biomédicos, mejorando el entendimiento de términos técnicos y estructuras específicas del dominio.
3. **Adaptabilidad:** Permite ajuste fino para tareas específicas como la clasificación de relaciones, lo que mejora significativamente la precisión.

4. **Eficiencia de recursos:** Aunque es computacionalmente intensivo, el preentrenamiento acelera la implementación práctica.

### **Limitaciones:**

- **Recurso intensivo:** Requiere hardware potente (GPUs o TPUs) para entrenamiento y ajuste.
- **Tamaño del modelo:** El ajuste puede ser difícil en proyectos pequeños o con datos limitados.

### **Justificación para uso:**

- BERT y sus variantes son ideales para proyectos biomédicos, ya que entienden el contexto técnico y manejan relaciones complejas. Si los abstracts son largos o las relaciones son altamente contextuales, un Transformer ajustado al dominio (como **BioBERT**) es la mejor opción.

## **Recomendación para el Proyecto de Clasificación de Relaciones Biomédicas**

- **¿Cuándo usar RNN?**
  - Si tienes recursos limitados, datos pequeños y las relaciones entre entidades dependen principalmente de palabras cercanas.
- **¿Cuándo usar LSTM?**
  - Si las relaciones son más complejas o están separadas por distancias moderadas dentro del abstract. También si el proyecto requiere equilibrio entre capacidad y costo computacional.
- **¿Cuándo usar BERT o Transformers?**
  - Si tienes acceso a recursos computacionales, y las relaciones dependen de contexto distante o altamente técnico. Las variantes preentrenadas como **BioBERT** están optimizadas para datos biomédicos y proporcionan los mejores resultados en clasificación avanzada.}

# Referencias

## Comparación de Modelos de Secuencia: RNN, LSTM, GRU y Transformers

- AIML. (2023). *Compare the different Sequence models (RNN, LSTM, GRU, and Transformers)*. Recuperado de [AIML](#)

## BertSRC: Clasificación de Relaciones Semánticas Basada en Transformers

- Lee, Y., Son, J., & Song, M. (2022). BertSRC: transformer-based semantic relation classification. *BMC Medical Informatics and Decision Making*, 22, Artículo 234. Recuperado de [BMC Med Info & Decision Making](#)

## Bioformer: Un Modelo de Lenguaje Transformer Eficiente para Minería de Textos Biomédicos

- Fang, L., Chen, Q., Wei, C.-H., Lu, Z., & Wang, K. (2023). Bioformer: an efficient transformer language model for biomedical text mining. *arXiv preprint arXiv:2302.01588*. Recuperado de [arXiv](#)

## Explorando Métodos de Aprendizaje Profundo para el Reconocimiento de Enfermedades Raras y sus Manifestaciones Clínicas en Textos

- Segura-Bedmar, I., Camino-Perdonas, D., & Guerrero-Aspizua, S. (2021). Exploring deep learning methods for recognizing rare diseases and their clinical manifestations from texts. *arXiv preprint arXiv:2109.00343*. Recuperado de [arXiv](#)

## BERT-GT: Extracción de Relaciones n-arias entre Oraciones con BERT y Graph Transformer

- Lai, P.-T., & Lu, Z. (2021). BERT-GT: Cross-sentence n-ary relation extraction with BERT and Graph Transformer. *arXiv preprint arXiv:2101.04158*. Recuperado de