



1. DATOS BÁSICOS DEL TFG:

Título: Transformers y el mecanismo de atención

Descripción general (resumen y metodología):

Las redes neuronales son estimadores universales de funciones y han experimentado un gran avance en capacidad en los últimos años debido a su flexibilidad y el incremento de la capacidad computacional disponible, permitiendo usos cada vez más complejos. Ejemplos recientes incluyen el desarrollo de modelos denominados "Transformers" que son capaces de trabajar con datos multimodales y son la punta de lanza de la IA generativa.

Este TFG explora y compara diferentes arquitecturas de Transformers y mecanismos de atención, la base de los Transformers, desde una perspectiva matemática, con el objetivo de explorar y describir su funcionamiento y cómo se han adaptado en su uso a diferentes tipologías de datos.

La metodología de este trabajo se basa en la búsqueda bibliográfica exhaustiva y la investigación experimental. A continuación, se detalla cada uno de estos aspectos:

- **Selección de Fuentes:** Se realizará una búsqueda exhaustiva de literatura científica en bases de datos reconocidas como Scopus, WoS, IEEE Xplorer, Springer, Google Scholar y arXiv. Se seleccionarán artículos, conferencias, tesis y libros relevantes que aborden tanto las arquitecturas más prominentes de Transformers como los mecanismos de atención.
- **Revisión Sistemática:** Se llevará a cabo una revisión sistemática de la literatura para identificar, evaluar y sintetizar los hallazgos más relevantes sobre la temática del TFG. Se prestará especial atención a estudios comparativos y revisiones que analicen múltiples mecanismos de atención.
- **Clasificación y Análisis:** Los estudios encontrados se clasificarán según las arquitecturas de Transformers utilizadas y se analizarán en términos de su aplicabilidad, efectividad y limitaciones. Se elaborará una tabla comparativa para resumir las características clave de cada técnica, incluyendo métricas de desempeño y áreas de aplicación.

Tipología: Estudio de casos, teóricos o prácticos, relacionados con la temática del Grado.

Objetivos planteados:

- Explorar los fundamentos teóricos de las redes neuronales (en particular, modelos de Transformers), así como describir sus propiedades matemáticas.
- Estudiar los fundamentos teóricos del mecanismo de atención. Diferencias en los diferentes tipos de datos.
- Explorar arquitecturas modernas de Transformers y su aplicación.

Bibliografía básica:

- [1] Christopher M. Bishop. Pattern Recognition and Machine Learning. (Springer, 2006)
- [2] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, (MIT Press, 2016). Available at <http://www.deeplearningbook.org>.
- [3] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).
- [4] Lin, Tianyang, et al. "A survey of transformers." AI open 3 (2022): 111-132.

[5] Khan, Salman, et al. "Transformers in vision: A survey." ACM computing surveys (CSUR) 54.10s (2022): 1-41.

Recomendaciones y orientaciones para el estudiante:

Se recomienda al alumno haber cursado las siguientes asignaturas:

Inferencia estadística, Análisis Matemático, Modelos matemáticos II, Métodos Numéricos.

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: JULIÁN LUENGO MARTÍN

Ámbito de conocimiento/Departamento: CIENCIAS DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL

Correo electrónico: julianlm@ugr.es

3. COTUTOR/A DE LA UGR (en su caso):

Nombre y apellidos:

Ámbito de conocimiento/Departamento:

Correo electrónico:

4. COTUTOR/A EXTERNO/A (en su caso):

Nombre y apellidos:

Correo electrónico:

Nombre de la empresa o institución:

Dirección postal:

Puesto del tutor en la empresa o institución:

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos: Carmen Vibora Zafra

Correo electrónico: carmenvibora@correo.ugr.es