



1. DATOS BÁSICOS DEL TFG:

Título: Evaluación de Políticas de Entrenamiento de LLMs con Estructuras GNN

Descripción general (resumen y metodología):

1.1. Motivación y contexto

Los Modelos de Lenguaje de Gran Escala (Large Language Models, LLMs) han revolucionado el campo del procesamiento del lenguaje natural. Modelos como GPT-4, LLaMA 2 o Gemini han demostrado capacidades emergentes sorprendentes, pero su entrenamiento sigue siendo un proceso costoso y empíricamente guiado. Una de las fases más críticas es la definición de la **política de entrenamiento**: la secuencia de decisiones sobre tasas de aprendizaje, currículos de datos, estrategias de fine-tuning y mecanismos de retroalimentación como RLHF y DPO.

Paralelamente, las Redes Neuronales de Grafos (GNNs) han emergido como paradigma poderoso para modelar estructuras relacionales complejas. Su capacidad para propagar información a través de nodos los hace especialmente adecuadas para representar dependencias entre hiperparámetros y decisiones de entrenamiento.

Trabajos recientes como MuseGraph demuestran que es posible integrar GNNs y LLMs de forma genérica para minería de grafos, logrando rendimiento estado-del-arte en clasificación de nodos, predicción de enlaces y generación de texto a partir de grafos. Más recientemente, SIT-Graph ha extendido este paradigma con ajuste de instrucciones semisupervisado, obteniendo mejoras de más del 20% en escenarios con escasez de etiquetas. Estos avances motivan directamente la investigación propuesta: si GNNs pueden modelar efectivamente la estructura de grafos para minería genérica, pueden igualmente modelar el grafo de decisiones que constituye una política de entrenamiento de LLMs.

Adicionalmente, el método PGMRA (Policy Gradient for Multi-objective Reward Alignment) ofrece un marco teórico relevante para evaluar y optimizar políticas de entrenamiento bajo múltiples objetivos de recompensa simultáneos, incluyendo calidad de generación, alineación con preferencias humanas y eficiencia computacional.

1.2. Descripción del problema

El entrenamiento de un LLM puede conceptualizarse como un grafo de decisiones: cada nodo representa un estado del modelo o una decisión de entrenamiento (elección del optimizador, mezcla de datos, temperatura de muestreo), y cada arista codifica la dependencia causal o temporal entre decisiones. Evaluar qué política de entrenamiento conduce a mejores resultados es un problema combinatorio de alta dimensión.

MuseGraph muestra que la representación textual compacta de grafos mediante descripciones de ego-grafos, vecinos de un salto y paseos aleatorios —junto con el ajuste de instrucciones con Chain-of-Thought (CoT) generadas por GPT— permite a los LLMs capturar estructura semántica y topológica simultáneamente. Este marco de descripción compacta de grafos, adaptado al dominio de políticas de entrenamiento, constituye la base metodológica de este TFG.

Por su parte, SIT-Graph resuelve el cuello de botella de la escasez de etiquetas mediante un pipeline iterativo de entrenamiento semi-supervisado: el modelo genera pseudo-respuestas filtradas por confianza para nodos no Propuesta TFG. Curso 2026/2027 GRADO: Grado en

Estadística CÓDIGO DEL TFG: etiquetados y los incorpora progresivamente al ajuste fino. Este mecanismo es directamente aplicable al contexto de evaluación de políticas de entrenamiento, donde la generación de datos etiquetados es costosa computacionalmente.

Finalmente, PGMRA proporciona el componente de optimización: dado un conjunto de políticas evaluadas por la GNN, el gradiente de política permite actualizar hacia configuraciones que maximicen un escalar de recompensa multiobjetivo, evitando búsquedas exhaustivas en el espacio de hiperparámetros.

1.3. Alcance del trabajo

El trabajo se enmarca en un TFG de investigación aplicada con las siguientes características:

- Se trabajará con modelos LLM de escala reducida (GPT, TinyLlama, LLaMA) para garantizar viabilidad computacional.
- Se adoptará el marco de descripción compacta de grafos de MuseGraph para representar políticas de entrenamiento como grafos de ego con nodos de hiperparámetros, etapas y conjuntos de datos.
- Se implementará un pipeline semi-supervisado inspirado en SIT-Graph para reducir la dependencia de experimentos de entrenamiento completamente etiquetados.
- Se evaluarán arquitecturas GNN (GCN , GAT , GraphSAGE) como evaluadores de calidad de política, con supervisión basada en perplejidad, MMLU y HellaSwag.
- Opcionalmente, se integrará PGMRA como módulo de optimización de políticas sobre el espacio de grafos evaluado.

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

Objetivos planteados:

Objetivo general

Diseñar, implementar y evaluar un framework basado en redes neuronales de grafos capaz de modelar y puntuar políticas de entrenamiento de LLMs, integrando técnicas de descripción compacta de grafos , ajuste semi-supervisado de instrucciones y optimización multiobjetivo de políticas mediante gradiente , con el fin de identificar configuraciones de entrenamiento más eficientes y de mayor calidad.

Objetivos específicos

- O1 Revisión bibliográfica Revisar el estado del arte en LLMs (Transformer, RLHF, DPO) y GNNs (GCN, GAT, GraphSAGE), con especial énfasis en los marcos MuseGraph, SIT-Graph y PGMRA para identificar brechas y oportunidades.
- O2 Formalización del grafo de políticas Definir formalmente la representación en grafo de una política de entrenamiento adoptando la ontología de nodos (hiperparámetros, etapas, datasets), tipos de aristas y atributos del marco MuseGraph, adaptada al dominio de entrenamiento de LLMs.
- O3 Descripción compacta de políticas Implementar el mecanismo de generación adaptativa de entradas (node energy, vecinos clave, paseos aleatorios) de MuseGraph para producir descripciones textuales compactas de políticas de entrenamiento dentro de los límites de tokens de los LLMs.
- O4 Construcción del dataset (semi-supervisado) Generar un dataset de políticas etiquetadas mediante experimentos controlados con LLMs de escala reducida. Aplicar el pipeline semi-supervisado de SIT-Graph con pseudo-respuestas filtradas por confianza para ampliar el dataset con nodos no etiquetados.

- O5 Diseño e implementación GNN Implementar y comparar GCN, GAT y GraphSAGE como evaluadores de políticas, utilizando PyTorch Geometric, con ajuste de instrucciones CoT generadas mediante el mecanismo de MuseGraph.
- O6 Evaluación experimental Diseñar un protocolo de evaluación que mida capacidad predictiva de las GNNs, correlación con métricas reales (perplejidad, MMLU, HellaSwag) y mejora obtenida por el pipeline semisupervisado respecto a la línea base supervisada.
- O7 Análisis de interpretabilidad Aplicar GNNExplainer y attention rollout sobre los modelos GAT para identificar qué componentes de la política tienen mayor impacto en el rendimiento final del LLM.
- O8 Optimización PGMRA (opcional) Integrar PGMRA como función de valor para optimizar políticas de entrenamiento mediante gradiente de política sobre el grafo evaluado, buscando configuraciones de Paretoóptimas bajo múltiples objetivos (rendimiento, alineación, coste).
- O9 Documentación y reproducibilidad Publicar código, dataset y resultados en repositorio abierto con documentación completa, garantizando la reproducibilidad de todos los experimentos.

Bibliografía básica:

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. & Polosukhin, I. Attention is all you need. *Advances in Neural Information Processing Systems* 30 (2017).
- Arnedo, J., del Val, C., de Erausquin, G. A. et al. PGMRA: a web server for (phenotype × genotype) many-to-many relation analysis in GWAS. *Nucleic Acids Research* 41, W142-W149 (2013).
- Zwir, I., Arnedo, J., Del-Val, C. et al. Uncovering the complex genetics of human personality. *Molecular Psychiatry* 25, 2295-2312 (2020).
- Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations* (2017).
- Schlichtkrull, M., Kipf, T. N., Bloem, P. et al. Modeling relational data with graph convolutional networks. *European Semantic Web Conference*, 593-607 (2018).
- Hamilton, W. L., Ying, R. & Leskovec, J. Inductive representation learning on large graphs. *Advances in Neural Information Processing Systems* 30 (2017).
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P. & Bengio, Y. Graph attention networks. *International Conference on Learning Representations* (2018).
- Ying, R., Bourgeois, D., You, J., Zitnik, M. & Leskovec, J. GNNExplainer: Generating explanations for graph neural networks. *Advances in Neural Information Processing Systems* 32 (2019).
- Yao, L., Mao, C. & Luo, Y. Graph convolutional networks for text classification. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 7370-7377 (2019).
- Tan, Y., Lv, H., Huang, X., Zhang, J., Wang, S. & Yang, C. MuseGraph: graph-oriented instruction tuning of large language models for generic graph mining. *arXiv 2403.04780* (2024).
- Song, Z. & King, I. Semi-supervised instruction tuning for large language models on text-attributed graphs. *Companion Proceedings of the ACM Web Conference 2026 (WWW '26)* (2026).

Recomendaciones y orientaciones para el estudiante:

Plan de trabajo

- Fundamentos y revisión Revisión bibliográfica.
- Estudio de frameworks (PyTorch, HuggingFace, PyTorch Geometric).
- Definición del alcance concreto.
- Diseño de representación Formalización del grafo de políticas con la ontología de MuseGraph. Implementación del módulo de descripción compacta (node energy, vecinos, paseos aleatorios).
- Generación de datos Experimentos de entrenamiento con LLMs reducidos para generar el dataset etiquetado base. Aplicación del pipeline semi-supervisado de SIT-Graph para ampliar con nodos no etiquetados.
- Implementación GNN + instrucción tuning Desarrollo de modelos GNN evaluadores con ajuste de instrucciones CoT estilo MuseGraph. Comparativa GCN vs GAT vs GraphSAGE.

- Evaluación, interpretabilidad y PGMRA Evaluación experimental completa. Análisis con GNNExplainer. Integración opcional de PGMRA para optimización de políticas.
- Redacción y entrega Escritura de la memoria. Preparación de la presentación oral. Publicación del repositorio.

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: JORGE SERGIO IGOR ZWIR NAWROCKI

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

Correo electrónico: igorzwir@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:

Centro de convenio Externo: Washington University

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos:

Correo electrónico: