

# ***Fine-tuning e Análise do Aprendizado de Modelos de Linguagem Navegando por Labirintos em ASCII***

**Anderson N. Silva<sup>1</sup>**

<sup>1</sup>Departamento Acadêmico de Informática (DAINF) –  
Universidade Tecnológica Federal do Paraná (UTFPR)  
Av. Sete de Setembro, 3165 – 80230-901 – Curitiba – PR – Brasil

`andersonnogueira@alunos.utfpr.edu.br`

**Abstract.** *10 lines and must be in the first page of the paper.*

**Resumo.** *Este meta-artigo descreve o estilo a ser usado na confecção de artigos e*

## **1. Introdução**

## **2. Fundamentos e Trabalhos Relacionados**

Esta seção discute os trabalhos relacionados e os fundamentos teóricos principais relacionados com o tema proposto.

### **2.1. LLMs - Large Language Models**

Os modelos de Linguagem de Grande Escala (*Large Language Models - LLMs*) são arquiteturas computacionais baseadas em redes neurais profundas, predominantemente do tipo *transformer*, desenvolvidos para processar e gerar linguagem natural. Esses modelos são treinados com grandes volumes de dados textuais, com os quais conseguem (computar/aprender) padrões estatísticos da linguagem humana, tornando-os capazes de realizar uma ampla gama de tarefas, grande parte sem treinamento supervisionado específico (e.g. sumarização, resposta a perguntas e tradução). A característica comumente associada ao tamanho desses modelos é uma medida em bilhões de parâmetros que está diretamente associado à sua capacidade de adaptação a contextos diversos e sua generalização. A arquitetura *transformer*, introduzida por [Vaswani et al. 2023], faz parte da estrutura central da maioria dos LLMs atuais, e destaca-se pelo uso de mecanismos de atenção que possibilitam a modelagem de dependências contextuais de longo alcance.

### **2.2. Fine-tuning**

O *fine-tuning* é uma técnica utilizada para adaptar *LLMs* a tarefas específicas, por meio de re-treinamento de parte ou totalidade de seus parâmetros com conjuntos de dados mais específicos. Em seguida da etapa inicial de pré-treinamento, na qual um *LLM* aprende as representações gerais linguagem a partir de grandes corpora, o *fine-tuning* permite especializar o comportamento do modelo para contextos mais delimitados como, por exemplo, geração de código, classificação de sentimentos ou, como no presente trabalho, navegação por labirinto em caracteres ASCII. Existem diferentes estratégias de *fine-tuning*:

- Supervised Fine-Tuning (SFT):
- Direct Preference Optimisation (DPO):
- Reinforcement Learning from Human Feedback (RLHF):
- Odds Ratio Preference Optimization (ORPO):

### **2.3. Explicabilidade em LLMs**

Apesar de serem bem capazes em tarefas complexas, os LLMs são questionados com bastante frequência sobre seu caráter de “caixa-preta”

### **References**

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2023). Attention is all you need.