

Elton Cardoso do Nascimento 233840

IA024 - Redes Neurais Profundas para Processamento de Linguagem Natural - 1s2024

# Leitura do Artigo "Language Models are Unsupervised Multitask Learners"

---

A principal contribuição do artigo foi demonstrar como um modelo de linguagem pode generalizar para diferentes tarefas sem treino prévio nesta tarefa a medida que a escala do modelo aumenta. O modelo torna-se então capaz de realizar outras tarefas condicionando-o apenas em sua própria entrada, indicando qual a tarefa que precisa ser realizada.

Além dos modelos treinados, em especial o "GPT-2", uma outra contribuição relevante é a metodologia e criação do dataset "WebText", criado a partir de textos na web filtrados por qualidade utilizando links tirados do Reddit. Ainda em relação aos datasets, o trabalho mostra como o modelo é capaz de generalizar para outros datasets nunca vistos, tendo uma outra contribuição a criação de um filtro capaz de calcular a sobreposição entre datasets.

É importante destacar também como o trabalho contribuiu de forma geral para o avanço da área e para a revolução que estamos tendo no uso de LLMs hoje.