Day 23: Few-Shot & Zero-Shot Learning

Introduction

Few-shot Learning

What?

- Powerful techniques in tasks involving natural language processing (NLP) and computer vision
- Sometimes acquiring large amounts of labeled data is impractical or expensive
- Enable models to generalize to new tasks with minimal or no labeled data
- Few-shot and zero-shot learning allow systems to perform tasks with minimal or no task-specific training data.
- In traditional ML, large amounts of labeled data are required to train models.
- These approaches have become increasingly popular, especially with the development of large language models like GPT-3
- Ability to generalize from a small number of examples (typically 1-5) provided during inference
- And then tasked with performing the same operation on new, unseen data.
- Unlike traditional ML, few-shot learning can generalize from just a small set of examples
- EG: Real-world scenarios such as medical diagnoses, where labeled data is scarce

How does it work?

- Meta-learning (also known as "learning to learn") involves training models that can quickly adapt to new tasks
- Exposed to many different tasks during training, allowing the model to learn how to learn from a small number of examples.
- Models like GPT-3 use a few-shot approach by adding several labeled examples in the prompt to help the model understand how to classify or complete a task.

Applications

- Sentiment Analysis
- Object Detection
- Language Translation

Zero-shot learning

- Goes one step further
- Ability to generalize to tasks it has never seen before, without being provided with any training examples
- By leveraging pre-trained models that can generalize across tasks based on their understanding of language or visual features.
- Can be achieved using models like GPT-3 by asking the model to perform a task without providing any labeled examples in the prompt.
- relies on its general knowledge and understanding of language

How does it work?

- Relies on powerful pre-trained models that have been trained on vast amounts of general data, such as large language models or vision models
- learn to encode generalized knowledge that can be applied to new tasks
- Model trained on diverse text data can understand concepts like "positive sentiment" or "negative sentiment" without being specifically trained for sentiment analysis
- requires models to have a deep understanding of both language and concepts

Applications

- Text Classification
- Intent Recognition
- Image Classification

How GPT-3 Performs Few-shot and Zero-shot Learning

GPT-3, the third iteration of OpenAI's Generative Pre-trained Transformer, is a powerful language model capable of performing few-shot and zero-shot learning.

- vast size (175 billion parameters)
- extensive pre-training on a diverse corpus of text data
- generalize across various tasks, even with little or no task-specific training.

```
Review: "The product was great!"
Sentiment: Positive

Review: "I did not like the product at all."
Sentiment: Negative

Review: "The delivery was fast and the product quality is amazing!"
Sentiment:
```

- Even though GPT-3 was not specifically trained on this task, it leverages its general understanding of language and the context from the few examples to predict the sentiment of the new review

```
Classify the following review as Positive or Negative:
"The delivery was slow and the product was broken."
```

- Zero-shot

Hands-on Project: GPT-3 for Zero-shot Text Classification