SOCIETAL ATTITUDES TOWARD AI: A TEMPORAL STUDY

PROJECT REPORT

Amittai Siavava
Dartmouth College
amittai.j.wekesa.24@dartmouth.edu

Aimen Abdulaziz
Dartmouth College
aimen.a.abdulaziz.25@dartmouth.edu

Angelic McPherson
Dartmouth College
angelic.mcpherson.25@dartmouth.edu

November 22, 2023

ABSTRACT

Artificial Intelligence (AI) is a hot topic, especially recently. This uptick in public interest in AI is driven by the improved capabilities of deep learning architectures such transformers. The emergent capabilities of such models, including GPT [1], has immense utility in other fields. So why has this dawn of AI has been met with conflicted sentiments?

The very fact that recent neural network–based models have immense capabilities leaves professionals anxious about their jobs and creative products. The ethics of deploying AI models in critical, user-facing applications where errorneous predictions have real-world consequences is also a concern. In fact, a faction of industry leaders has even called for AI experiments to be paused until proper regulations are in place because of the risk of runaway AI.

There is also a lingering question around the ownership of work generated by AI models. When a model's output is clearly derived from a copyrighted artifact, who owns the output? The original creator of the imitated artifact may not want to be associated with the new imitation, since it may have inferior quality or even be offensive. The same concern applies to the creators of the AI models themselves. At the same time, imitations of copyrighted works *without permission* have their own legal implications.

This research project studies the societal attitudes toward AI, both currently and how they have evolved over the years, as a way to understand how different events have shaped the public's perception of AI. We use topic modeling, sentiment analysis, and procrustes analysis to analyze relationships across time periods and extract insight into the changing story of artificial intelligence.

Keywords AI · Machine Learning · Ethics

1 Introduction

We sought to find a *representative* yet *accurate* sample of the public's opinion on AI. We considered multiple potential sources of data, and their tradeoffs:

  (i) *Research papers* are the most cutting-edge and factually correct, but they tend to dive into detailed exposition of novel model architectures and ideas, which is disconnected from the public's opinion.

  (ii) *Social media posts* are most accessible to the public, yet they are often too short to offer a nuanced opinionand are not fact-checked so they are prone to inaccuracies.

(iii) *News articles* can be sensationalized and biased, but they are often longer (than social-media posts), fact-checked, and backed by current events and innovations. This keeps them (or at least the good ones) in touch with both the public's sentiments about AI/technology and in touch with new innovations.

We decided to use news articles as our target data source, since they are a good compromise. However, we eventually limited our domain to a few specific news sources that are both reliable and well-known, as explained in **??**. We then considered potential forms of analysis to use as a lens to study the data:

(i) *Topic modeling* can be used to identify the most common topics in the data. We can then focus on these topics and compare them across years or in individual years using methods such as procrusties analysis. It can also be insightful to see the most prominent topics in conversations in given periods, or how specific topics such as *ethics* became more or less emphasized after certain events, such as national elections.

(ii) *Sentiment analysis* can be used to identify sentiments or tones toward AI and how they change. This can be insightful in identifying when the public's attitude toward AI became more positive or negative.

(iii) *Procrustes analysis* can be used to identify shifts in conversation in specific time periods, and help highlight periods of particular interest.

References

[1]   Alec Radford et al. "Language models are unsupervised multitask learners". In: (2019).