Extreme Over-Representation Bias for Understanding Individuals and Small Groups

How extreme over representation can augment expert analysts

April 2025

By Dwayne Phillips

Summary

Representation bias plagues the building of large language models (LLMs). There is, however, an opportunity that may come from representation bias in fine-tuning an LLM. Extreme over representation of individuals and small groups can augment expert analysts and provide greater understanding and far fewer misunderstandings of other groups and other individuals. This essay concludes with a report on a demonstration of the technique.

Fine Tuning a Large Language Model and Bias

In deep learning, fine-tuning is an approach to transfer learning in which the parameters of a pre-trained neural network model are trained on new data. (Wikipedia Contributors, 2023 December 12)

Fine-tuning an LLM is a basic step in using LLMs and other types of AI. The simple idea is to slant an LLM towards a particular endeavor so that it uses the vocabulary of that endeavor. For example, theologians use “grace” in one way while ballet masters use “grace” differently. Fine tuning LLMs adds the necessary context.

One problem that can occur in fine-tuning an LLM is representation bias. This is a type of sampling bias. In statistics, sampling bias is a bias in which a sample is collected in such a way that some members of the intended population have a lower or higher sampling probability than others. It results in a biased sample. (Wikipedia Contributors, 2023 December 11)

There is, however, an opportunity that may come from representation bias in fine-tuning an LLM.

This sampling or representation bias easily creeps into LLMs and their fine tuning. The drawbacks of representation bias are obvious.

Consider a simple example. Persons are either "right-handed" or "left-handed." Each "handed-ness" case should represent 50% of the training samples. Wrong. That would over-represent the left-handed persons and cause companies to build far too many tools specific to left-handed persons. This simple example would cause an economic disaster for a company.

Extreme Over-Representation Bias

There is, however, an opportunity that may come from representation bias in fine-tuning an LLM. This proposes fine-tuning an LLM using only new data from one individual. Henceforth, "individual" may be read as "one person or a small number of persons who are like minded."

An LLM fine-tuned on the data of an individual would respond far more like that individual than an LLM fine-tuned on another endeavor. Consider fine-tuning on just Martha Stewart as opposed to all fashion models of the 20th century.

Oh, Now I Understand

An LLM fine-tuned using only new data from one individual may promote understanding of that individual. Consider the prompt, "Characterize the individual." The LLM would provide a description of the individual based on all the individual's data used in fine tuning.

There are many other prompts that can help understand the individual. Consider:

- What thoughts on foreign trade does the individual express in their data?
- Where is the individual likely to go on vacation?
- What type of good and services is the individual likely to purchase?

An LLM fine-tuned using only new data from an individual can provide answers to these and many other questions that seek understanding. Those answers, while not guaranteed correct, will provide insight into the individual.

Note: there are nefarious uses for employing extreme over-representation bias. There are nefarious uses for just about any AI or just about any tool of any type. This essay will not describe those nefarious uses or how to combat them.

Augmenting the Human Expert Analyst

Understanding or seeking to understand of an individual has been attempted for centuries. Biographers study the data of individuals for decades. They are then asked, "What would Abraham Lincoln say to our current President?"

There are expert analysts who study individuals. They are also asked the questions about the individual to help understand what the individual would do or prefer. Such insights help industries prepare for the marketplace and shifts in preferences.

The goal of the human expert analyst is better understanding on an individual. That is also the goal of extreme over-representation bias.

There are many more individuals that we wish to understand than there are human expert analysts. Extreme over-representation bias augments the pool of human expert analysts.

An Example of Extreme Over-Representation Bias

Charles Haddon Spurgeon (19th June 1834 – 31st January 1892) was an English Particular Baptist preacher. Spurgeon was pastor of the congregation of the New Park Street Chapel (later the Metropolitan Tabernacle) in London for 38 years. (Wikipedia Contributors, 2023 December 26)

The sermons of Mr. Spurgeon were written and published for posterity. There are some 40,000 pages of these available online. This is an excellent source of data for fine-tuning an LLM.

NotebookLM is a service provided by Google. It helps synthesize and understand information from documents. In one use case, users load documents into NotebookLM to fine-tune its LLM on those documents. (This essay is not an endorsement of NotebookLM or any other systems from Google. It is used as one example that illustrates the technique of extreme over-representation bias.)

As an example of extreme over-representation bias, I loaded 40,000 pages of writings from Mr. Spurgeon into NotebookLM. I then prompted NotebookLM with questions like those listed earlier.

The system responded with detailed answers. It explained what Mr. Spurgeon might write in a letter to current world leaders and how it would address different world leaders in different settings differently. All the answers referred back to the principles found in the 40,000 pages of Mr. Spurgeon's writings.

A biographer or expert analyst of Mr. Spurgeon could provide such answers. I would first have to find a biographer or expert analyst and ask those questions. Limited resources would prevent such.

This example demonstrated some understanding of Mr. Spurgeon and the utility of extreme over-representation bias.

The sources of these 40,000 pages were:

https://www.monergism.com/thethreshold/sdg/spurgeon/The_Sermons_of_Charles_Spurgeon_-_C_H_Spurgeon.pdf

https://www.spurgeongems.org/spurgeon-sermons/

References

Wikipedia Contributors. (2023, December 11). Sampling bias. In Wikipedia. https://en.wikipedia.org/wiki/Sampling_bias

Wikipedia Contributors. (2023, December 12). Fine-tuning (deep learning). In Wikipedia. https://en.wikipedia.org/wiki/Fine-tuning_(deep_learning

Wikipedia Contributors. (2023, December 26). Charles Spurgeon. In Wikipedia. https://en.wikipedia.org/wiki/Charles_Spurgeon