Fine-tuning human for LLM projects

Rehan Guha rehanguha29@gmail.com
Data Scientist
Vodafone
India

Editor: Self on September 2023

Abstract

In recent years after emerging of LLM in the space of NLP, tools like ChatGPT, BARD, DALL-E, took over the market and daily lives. We will be focusing on the human like output capability from LLM's . There is a huge reservation of this technology due to multiple scenarios like security, accuracy, relevance, etc... In this paper, I will talk about a method which I have designed to fine tune a human being to lower the expectation from LLM outputs and increase the acceptance rate of the final product. This technique is more of a psychological method than a technological way to improve the models output to be more human like.

Keywords: NLP, Large Language Models, Psychology, Cognition Control

1. Introduction

When we have designed a tool using embedding models and LLM's (Touvron u. a.) on a specific domain use case. As a non-domain expert it was giving a decent result and the accuracy was high. Unfortunately when it went to the business team it was bluntly rejected saying it was not human like and this tool will add more of an overhead than help in their day-to-day solution (Farrell u. a.) (Shi u. a.) (Vaithilingam u. a.). After a long discussion with the business we found there are some below points which were bugging them:

- The output is not human like who knows the domain so, it is creating a mental barrier to accept it
- The accuracy of handful output for the prompts are not that high
- Some UX Design related optimization was required, etc...

In this paper we tacked the first problem where the business team thought the chatbot reply was not like a domain expert.

So to tackle that, I have designed a multiple choice questionnaire to adapt a human expectation, but it will be a type of cognition control strategy.

The way I designed the MCQs here, it will slowly indicate the human user (who is interacting with test) about the right expectation from the model using multiple scenarios and positive feedback.

The concept which I thought worked by adjusting the expectation of the end user by promoting what can you expect from the LLM models. Testing LLM models are quite challenging as there is no standard framework around it. This actually worked for two use cases where I have implemented this. It is very complicated to verify and evaluate the efficiency of the method proposed method whether it will work properly or not but after practically implementing this I can conform it worked seamlessly for 2 use cases.

2. Background & Related Work

Cognitive control (Farrell u. a.) (Luna u. a.) (Musslick und Cohen) is a broad class of mental operations that include goal representation, attention allocation, and stimulus-response mapping. Cognitive control is important when there is competition for limited mental resources. It helps reduce uncertainty in decision-making by controlling what information reaches focused awareness. According to the Expected Value of Control theory, people integrate information about the expected reward and efficacy of task performance to determine the expected value of control. They then adjust their control allocation (i.e., mental effort) accordingly (Frömer u. a.).

3. MCQ Template Used

I have used this below template for 2 projects which worked well and business team now understands the expectation from the chatbot. Below is the used template:

- LLM for Scenario X
  - Q(1)
    - Human Answer1
    - Human Answer2
    - LLM output
    - Human Answer3
  - Q(1)
    - Human Answer1
    - LLM output
    - Human Answer2
    - Human Answer3
  - Q(n/2)
    - Human Answer1
    - LLM output1
    - LLM output2
    - Human Answer2
  - Q(n)
    - Human Answer1
    - LLM output1
    - LLM output2

        * LLM output3

- LLM for Scenario X+1
    - ...
- LLM for Scenario X+n
    - ...

As the questions increases then no. of LLM outputs should increase which increase the probability of selection of LLM output choices this will boost the confidence and morale of the end user that they are able to guess the right source of the data.

4. Implementation

I have used the above template to design the MCQs for multiple scenarios. The way the feedback is designed it will help the users to know about the LLM outputs more in depth.

If the user selects Human answer then we will be showing the end-user a positive message saying "You are close enough but, the machine will provide you with the following answer …". If the end-user selects the LLM answer then we will again prompt with a positive message saying "Machine is thinking like you, and you are right! This is the answer which machine will generate."

We had around 20 questions for multiple set of documents more than 4 scenarios, and we tested the end users reaction before and after taking the test. The entire outlook about LLM models changed after going through the designed tests. Along with this it was informed that this is more of a preview and LLM will work as an aid and with continuous expert feedback and finetuning (Bill und Eriksson) (Ziegler u. a.) (Stiennon u. a.) we will be able to reach the desired accuracy with iterative improvements.

5. Conclusion

After going through multiple iterations we have seen some changes to the end users. This experiment was conducted for 2 projects and both projects was successful.

Some visible changes which I have observed are:

- Expectation got lowered
- Gave more constructive feedback (before the test the users were not that ready to accept the solution)
- Still had reservations but open to accept the future of the product

As it was informed to the end users that this product is more of a live preview and with human feedback this can be fine-tuned to perfection. This process is not a one-day process it takes time and with time this will surely pass the subject matter expert turing test or Feigenbaum test (Feigenbaum).

6. Future Work

There can be some more work to round up this method. This method can be more refined and some A/B acceptance testing can be performed on top of this to have a more strong validation. As I have used this in a corporate environment I was a tough for me to go through this test.

Along with the above and good feedback loop (Lee u. a.) (Gulcehre u. a.) (Zhang u. a.) (Bai u. a.) can be designed to solve the second have of the problem then it will be like to solve a single problem from multiple ends.

References

[Bai u. a.] Bai, Yuntao ; Kadavath, Saurav ; Kundu, Sandipan ; Askell, Amanda ; Kernion, Jackson ; Jones, Andy ; Chen, Anna ; Goldie, Anna ; Mirhoseini, Azalia ; McKinnon, Cameron ; Chen, Carol ; Olsson, Catherine ; Olah, Christopher ; Hernandez, Danny ; Drain, Dawn ; Ganguli, Deep ; Li, Dustin ; Tran-Johnson, Eli ; Perez, Ethan ; Kerr, Jamie ; Mueller, Jared ; Ladish, Jeffrey ; Landau, Joshua ; Ndousse, Kamal ; Lukosuite, Kamile ; Lovitt, Liane ; Sellitto, Michael ; Elhage, Nelson ; Schiefer, Nicholas ; Mercado, Noemi ; DasSarma, Nova ; Lasenby, Robert ; Larson, Robin ; Ringer, Sam ; Johnston, Scott ; Kravec, Shauna ; Showk, Sheer E. ; Fort, Stanislav ; Lanham, Tamera ; Telleen-Lawton, Timothy ; Conerly, Tom ; Henighan, Tom ; Hume, Tristan ; Bowman, Samuel R. ; Hatfield-Dodds, Zac ; Mann, Ben ; Amodei, Dario ; Joseph, Nicholas ; McCandlish, Sam ; Brown, Tom ; Kaplan, Jared: *Constitutional AI: Harmlessness from AI Feedback*. – URL http://arxiv.org/abs/2212.08073. – Zugriffsdatum: 2023-09-12

[Bill und Eriksson] Bill, Desiree ; Eriksson, Theodor: FINE-TUNING A LLM USING REINFORCEMENT LEARNING FROM HUMAN FEEDBACK FOR A THERAPY CHATBOT APPLICATION.

[Farrell u. a.] Farrell, Wendy ; Bogodistov, Yevgen ; Mössenlechner, Claudia: Is Academic Integrity at Risk? Perceived Ethics and Technology Acceptance of ChatGPT. . – URL https://aisel.aisnet.org/amcis2023/sig_ed/sig_ed/24

[Feigenbaum] Feigenbaum, Edward A.: Some challenges and grand challenges for computational intelligence. 50, Nr. 1, S. 32–40. – URL https://doi.org/10.1145/602382.602400. – Zugriffsdatum: 2023-09-13. – ISSN 0004-5411

[Frömer u. a.] Frömer, R. ; Lin, H. ; Dean Wolf, C. K. ; Inzlicht, M. ; Shenhav, A.: Expectations of reward and efficacy guide cognitive control allocation. 12, Nr. 1, S. 1030. – ISSN 2041-1723

[Gulcehre u. a.] Gulcehre, Caglar ; Paine, Tom L. ; Srinivasan, Srivatsan ; Konyushkova, Ksenia ; Weerts, Lotte ; Sharma, Abhishek ; Siddhant, Aditya ; Ahern, Alex ; Wang, Miaosen ; Gu, Chenjie ; Macherey, Wolfgang ; Doucet, Arnaud ; Firat, Orhan ; Freitas, Nando de: *Reinforced Self-Training (ReST) for Language Modeling*. – URL http://arxiv.org/abs/2308.08998. – Zugriffsdatum: 2023-09-12

[Lee u. a.] Lee, Harrison ; Phatale, Samrat ; Mansoor, Hassan ; Lu, Kellie ; Mesnard, Thomas ; Bishop, Colton ; Carbune, Victor ; Rastogi, Abhinav: *RLAIF: Scaling Reinforcement Learning from Human Feedback with AI Feedback*. – URL http://arxiv.org/abs/2309.00267. – Zugriffsdatum: 2023-09-12

[Luna u. a.] Luna, Beatriz ; Paulsen, David J. ; Padmanabhan, Aarthi ; Geier, Charles: Cognitive Control and Motivation. 22, Nr. 2, S. 94–100. – URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4285389/. – Zugriffsdatum: 2023-09-08. – ISSN 0963-7214

[Musslick und Cohen] Musslick, Sebastian ; Cohen, Jonathan D.: Rationalizing constraints on the capacity for cognitive control. 25, Nr. 9, S. 757–775. – URL https://www.sciencedirect.com/science/article/pii/S1364661321001480. – Zugriffsdatum: 2023-09-08. – ISSN 1364-6613

[Shi u. a.] Shi, Freda ; Chen, Xinyun ; Misra, Kanishka ; Scales, Nathan ; Dohan, David ; Chi, Ed H. ; Schärli, Nathanael ; Zhou, Denny: Large Language Models Can Be Easily Distracted by Irrelevant Context. In: *Proceedings of the 40th International Conference on Machine Learning*, PMLR, S. 31210–31227. – URL https://proceedings.mlr.press/v202/shi23a.html. – Zugriffsdatum: 2023-09-08. – ISSN: 2640-3498

[Stiennon u. a.] Stiennon, Nisan ; Ouyang, Long ; Wu, Jeff ; Ziegler, Daniel M. ; Lowe, Ryan ; Voss, Chelsea ; Radford, Alec ; Amodei, Dario ; Christiano, Paul: *Learning to summarize from human feedback*. – URL http://arxiv.org/abs/2009.01325. – Zugriffsdatum: 2023-09-12

[Touvron u. a.] Touvron, Hugo ; Martin, Louis ; Stone, Kevin ; Albert, Peter ; Almahairi, Amjad ; Babaei, Yasmine ; Bashlykov, Nikolay ; Batra, Soumya ; Bhargava, Prajjwal ; Bhosale, Shruti ; Bikel, Dan ; Blecher, Lukas ; Ferrer, Cristian C. ; Chen, Moya ; Cucurull, Guillem ; Esiobu, David ; Fernandes, Jude ; Fu, Jeremy ; Fu, Wenyin ; Fuller, Brian ; Gao, Cynthia ; Goswami, Vedanuj ; Goyal, Naman ; Hartshorn, Anthony ; Hosseini, Saghar ; Hou, Rui ; Inan, Hakan ; Kardas, Marcin ; Kerkez, Viktor ; Khabsa, Madian ; Kloumann, Isabel ; Korenev, Artem ; Koura, Punit S. ; Lachaux, Marie-Anne ; Lavril, Thibaut ; Lee, Jenya ; Liskovich, Diana ; Lu, Yinghai ; Mao, Yuning ; Martinet, Xavier ; Mihaylov, Todor ; Mishra, Pushkar ; Molybog, Igor ; Nie, Yixin ; Poulton, Andrew ; Reizenstein, Jeremy ; Rungta, Rashi ; Saladi, Kalyan ; Schelten, Alan ; Silva, Ruan ; Smith, Eric M. ; Subramanian, Ranjan ; Tan, Xiaoqing E. ; Tang, Binh ; Taylor, Ross ; Williams, Adina ; Kuan, Jian X. ; Xu, Puxin ; Yan, Zheng ; Zarov, Iliyan ; Zhang, Yuchen ; Fan, Angela ; Kambadur, Melanie ; Narang, Sharan ; Rodriguez, Aurelien ; Stojnic, Robert ; Edunov, Sergey ; Scialom, Thomas: *Llama 2: Open Foundation and Fine-Tuned Chat Models*. – URL http://arxiv.org/abs/2307.09288. – Zugriffsdatum: 2023-09-12

[Vaithilingam u. a.] Vaithilingam, Priyan ; Zhang, Tianyi ; Glassman, Elena L.: Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by

Large Language Models. In: *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery (CHI EA '22), S. 1–7. – URL https://dl.acm.org/doi/10.1145/3491101.3519665. – Zugriffsdatum: 2023-09-12. – ISBN 978-1-4503-9156-6

[Zhang u. a.] Zhang, Tianjun ; Liu, Fangchen ; Wong, Justin ; Abbeel, Pieter ; Gonzalez, Joseph E.: *The Wisdom of Hindsight Makes Language Models Better Instruction Followers*. – URL http://arxiv.org/abs/2302.05206. – Zugriffsdatum: 2023-09-12

[Ziegler u. a.] Ziegler, Daniel M. ; Stiennon, Nisan ; Wu, Jeffrey ; Brown, Tom B. ; Radford, Alec ; Amodei, Dario ; Christiano, Paul ; Irving, Geoffrey: *Fine-Tuning Language Models from Human Preferences*. – URL http://arxiv.org/abs/1909.08593. – Zugriffsdatum: 2023-09-12