

# Project proposal

Tomer Shimshi 203200480 Amit Damri 313191785

March 2024

- 1 Project title: Application of LLM (llama-2) to answer Rebbe Questions in Hebrew**
- 2 Names and e-mails of the contributing students:**

Tomer shimshi tomershimshi@gmail.com Amit Damri amitdamri@tau.ac.il

## **3 A brief description of the proposed project:**

In the proposed project we will try to fine tune a LLM model (llama-2 or some other good open source LLM) to answer Rebbe questions in Hebrew. A Rebbe Question is question asked by a Jewish person to his Rebbe and the Rebbe needs to answer them according to Jewish tradition law the questions can be about anything. In our project we will try to fine tune a LLM on a QA dataset which we will comprise from the following Rebbe questions site: <https://www.yeshiva.org.il>, split the questions to a train and test datasets and evaluate the models performance using the test dataset. We will want to test if it is even possible for an LLM to answer such a complicated questions that a Rebbe spends years studying all the Jewish books in order to answer. We can also propose that if the model does not give descent results to try and first fine tune it on all the impotent Jewish books and after that do an additional fine tuning over the Rebbe QA dataset which we will build (all the relevant Jewish books can be found on this site: <https://mechon-mamre.org/indexhe.htm>) If this project will succeed it can help many remote orthodox Jews that do not have a frequent access to a near by Rebbe but do require his free assistant and it can potentially shift the balance of power in the entire orthodox Jewish community in Israel since everyone will have a free access to valid Rebbe. We know that LLMs can answer questions in Hebrew but it is not so trivial to answer Rebbe Question and not many people are qualified to answer them.

## **4 Our assumptions and requirements for the project:**

We assume that since we know LLMs can give expert like results in other fields (such as computer science) we assume that our goal of answering Rebbe question is indeed feasible. We will also need access to a GPU based compute power, by using some quantization techniques the fine tuning can be done with relatively low compute power (compared to the compute power first used when the main model was trained)

## **5 Compute requirements:**

We will need access Via ssh to a server containing a strong GPU in order to complete the fine tuning on the QA dataset in relatively short period of time