**Assignment**

**Chatbot as a Personal Trainer**

Student’s Name: Vachan Giriyapur

Course: Natural Language Processing

Course Code: DTSC422

Prof. Manoranjan Dash

### Introduction and Motivation:

In recent years, chatbots have become the mainstream for sources of information, guides on certain topics and more. Products that are chatbots such as ChatGPT, DeepSeek AI and many more have changed the ways we retrieve information and conduct research. The world is moving towards chatbots being a necessity.

Fitness is a domain which is receiving a lot of interest today, especially amongst younger generations. Although starting to workout is simple, it involves harming one’s body to a certain extent and comes with an amount of risk. Hence, it is advised to have a personal trainer when one starts going to gym to prevent injuries at an early stage.The issue with having a personal trainer and buying a gym membership is that they could cost a lot of money. What if a chatbot could be used as a personal trainer? That would make fitness more cost-effective and provide a personal trainer in one's hand. Being a person who regularly goes to the gym, I was not aware of the workout plan I should follow given health issues and goals that were part of my profile. Using the intelligence of a chatbot, I could start my journey in an efficient manner.

Therefore, the assignment is to create a fitness chatbot that can also act as a personal trainer. When one asks a fitness related question, they will receive advice from the chatbot that is trained on data created to resolve fitness related issues. The chatbot algorithm makes use of Retrieval-Augmented Generation (RAG) and pre-trained transformer models. The chatbot is designed in a manner to retrieve information related to fitness and provide them as answers, which are refined by a generative model.

The objective of the assignment is as follows:

* Create an interactive chatbot to resolve fitness related queries
* Retrieve relevant answers from the dataset with the use of sentence embeddings
* Make use of pre-trained transformer generative models to refine the answers provided by the chatbot

### Literature Review:

Chatbot for fitness management using IBM Watson by Sai Rugved Lola:

* Lola, Dhadavi, Wang and Zhu from the University of Maryland wrote a paper discussing the potential of using IBM Watson to develop a chatbot for fitness management, the chatbot proposed was to create diet plans , training plans and many more applications.
* The paper states that a chatbot makes use of natural language processing, natural language understanding and other technologies provided by IBM Cloud Watson. The intention is to use the chatbot as a solution to manage workout plans, diet plans and provide reminders and rewards to customers as motivation to exercise.
* The paper also talks about the impact of AI and machine learning in the fitness industry, with automation aiding processes and improving overall customer relationships.

Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks by Patrick Lewis:

* Lewis, Pereze and Piktus make a case for using Retrieval-Augmented Generation Models and their use in tasks of natural language processing. The authors state that large pre-trained language models struggle with knowledge-intensive tasks, accessing knowledge with accuracy and updating the world's knowledge. They argue that RAG models generate more specific, diverse and factual language than parametric only models.
* Two RAG formulations are explored: one uses the same retrieved passages for the whole sequence, and the other uses different passages per token. RAG models set the state-of-the-art on three open-domain QA tasks and generate more specific, diverse, and factual language than parametric-only models.
* RAG models achieve state-of-the-art results on open-domain QA tasks (Natural Questions, WebQuestions, TriviaQA, CuratedTrec). They outperform BART on MS-MARCO and Jeopardy question generation tasks, generating more factual and specific questions. On FEVER fact verification, RAG approaches state-of-the-art without needing retrieval supervision.

An Empirical Study on Factors Influencing Consumer Adoption Intention of an AI-Powered Chatbot for Health and Weight Management by Chin-Yuan Huang:

* A study by Huang and Yang, looked to investigate the factors that influence consumer adoption intention of an AI-powered chatbot for health and weight management using a model known as UTAUT2. The health chatbot used in the study runs an APP platform and utilises artificial intelligence for health consultation, diet tracking and exercise tracking.
* The results of the study found that the model explained 87% of variance in behavioral intention. They found that a user’s habit was the strongest factor in predictions, after that performance expectancy, social influence and network externality were the biggest factors.
* The implications of the study were that it provided insights to developers of AI powered apps to align user benefits and habits. To increase the number of users, network externality and innovation play key roles. Despite the benefits of the app, the study sample might not represent general populations and depends on behavior user by user.

### Methodology:

The dataset used is a custom made dataset that contains 20 fitness related questions and answers. The file is a structured YAML file in the form of a dictionary, which is required by the transformer models used in the code. The data preprocessing included converting the questions and answers into structure embeddings for efficient retrieval of questions and answers related.

The libraries used are transformers, sentence-transformers, PyTorch and YAML. The YAML package is used to load the data which is in a yaml format, the YAML package is normally used for structure data storage. The PyTorch package gives the user the ability to use tensor operations and deep learning operations, which is key for calculating cosine similarity and helps when it comes to embeddings.

The transformers package helps load and apply the pre-trained transformer model for the text generation and other natural language processing tasks, the package helps us to use the Facebook Blenderbot. The Facebook Blenderbot 400M distill is a natural language processing model implemented in a transformer library, which uses the python language.

The sentence-transformers package gives access to the all-MiniLM-L6-v2 pre-trained embedding models which converts text into vector representations for the similarity search. Sentence Transformer all-MiniLM-L6-v2 is specifically designed to generate numerical embeddings (vectors) representing the semantic meanings of the sentences, allowing for comparison and analyzing the similarity between different questions by converting them into vectors. The AutoTokenizer and AutoModel packages load the tokenizer and model from the pre-trained transformer architectures, the packages tokenizer user queries to make them compatible with the transformers.

The questions from the dataset are converted into vector embeddings so that the search for relevant questions is efficient, the tokenization is done using the all-MiniLM-L6-v2 pre-trained model. When the user asks the chatbot a question, the question is encoded and then the most similar questions found from within the dataset to extract the best possible answer from the dataset.

While the bot could return the answer as it is, the Facebook BlenderBot-400 M refines the answer to return as an output, the pre-trained model rephrases the answer to return in a more conversational manner, providing a more human touch to the answers instead of generic answers.The code includes a user input loop to continuously take inputs and respond accordingly. The loop provides the bot the ability for real time interaction, allowing the chatbot to be fully functional.

### Results and Evaluation:

The chatbot works at a satisfactory level, providing accurate responses to all the fitness related questions asked, the chatbot has provided the right answer to all the questions asked so far and often returns the correct answer. The chatbot also returns the right answer or a reasonable answer to questions that have the same meaning as the questions in the dataset but could be rephrased in different manner. For example, when the bot was asked: “lose weight how?” instead of the question in the dataset which is phrased as “How to lose weight effectively?” , the bot returned the correct answer which was: “To lose weight, maintain a calorie deficit by eating fewer calories than you burn. Combine strength training and cardio for best results.” This is a result of using cosine similarity to search and extract answers to return as an output and the overall effectiveness of retrieval and generative models. Additionally, the chatbot provides answers promptly and almost immediately with a few seconds delay in case the question is majorly rephrased.

While the chatbot works in a satisfactory manner, the chatbot still lacks in certain aspects. The dataset still limits the chatbot. The chatbot could provide answers to fitness related questions, but does not fully act as a personal trainer, the bot would not be able to create a fitness plan for a user. The chatbot also still lacks dynamically as it may not be able to answer questions that may require a more dynamic response. The responses are still very much dependent on the quality of the dataset that is being utilized in the code. Hence, there may be a need to add a level of intelligence to the bot.

### Future Work and Applications:

The references used in the paper and results of the assignment provide the potential of using chatbots in the fitness industry. Hence the following improvements may be applied in the future to enhance the effectiveness of the chatbot:

The first improvement would be to create a transformer model specifically designed for the fitness domain and answer fitness related questions. The current code makes use of pre-trained models for tokenization and returning answers and as a result, the chatbot is not trained on the dataset that is utilized in this particular assignment. The dataset will also be expanded to increase the overall knowledge base with more fitness related questions and answers. Ideally, a more advanced version of the chatbot should be able create workout plans and diet plans to provide a more well-rounded user experience. After implementing the above changes, the goal would be to integrate the product as web based or mobile chatbot to create accessibility. Providing accessibility to a greater audience.

### References:

1. Lola, S. R., Dhadvai, R., Wang, W., & Zhu, T. (2021). Chatbot for fitness management using IBM Watson. *arXiv preprint arXiv:2112.15167*.

<https://arxiv.org/pdf/2112.15167>

1. Huang, C. Y., Yang, M. C., & Huang, C. Y. (2021). An empirical study on factors influencing consumer adoption intention of an AI-powered chatbot for health and weight management. *International Journal of Performability Engineering*, *17*(5), 422.

<https://www.ijpe-online.com/www.ijpe-online.com/EN/10.23940/ijpe.21.05.p2.422432>

1. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, *33*, 9459-9474.

<https://proceedings.neurips.cc/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf>