**Phase-1: Literature survey & Data Acquisition**

**1. Problem definition:**

* 1. **What is the problem all about?**

Ans. Building a chatbot to answer the FAQs on mental health.

* 1. **Why is this an important problem to solve?**

1. People who need the treatment to get themselves in a healthy mental state, used to have some questions in their minds. These questions have to be answered by domain experts in order to make sure that the patient gets the right knowledge and trusts the treatment. Till a few years back, these kinds of questions were handled with the help of human consultants. But, now the industry is moving towards chatbots in order to resolve customer queries.
2. Chatbot resolves the customer queries instantly whereas the customer needs to wait in queue to talk to the human customer executive. Chatbot saves the waiting time of customers.
3. Handling customer queries using chatbot is cost effective as compared to when this is done through human customer executives.
4. Bot provides 24\*7 support without getting tired.

All these advantages of chatbots makes this an important problem to solve from customer as well as business point of view.

* 1. **Business/Real-world impact of solving this problem.**

1. People who need the treatment to get themselves in a healthy mental state, used to have some questions in their minds. These questions have to be answered by domain experts in order to make sure that the patient gets the right knowledge and trusts the treatment. Till a few years back, these kinds of questions were handled with the help of human consultants. But, now the industry is moving towards chatbots in order to resolve customer queries.
2. Chatbot resolves the customer queries instantly whereas the customer needs to wait in queue to talk to the human customer executive. Chatbot saves the waiting time of customers.
3. Handling customer queries using chatbot is cost effective as compared to when this is done through human customer executives.
4. Bot provides 24\*7 support without getting tired.
5. **Dataset:** 
   1. **Source of the dataset.**
6. Kaggle
   1. **Explanation of each feature and datapoint available.**
7. Total 98 data points are available.
8. There are 3 columns- Unique ID, Question, Answer
   1. **Data Size and any challenges you foresee to process it?**

Ans. Data Size is 160 KB which can be easily processed.

* 1. **Tools (Pandas, SQL, Spark etc) that you will use to process this data.**

Ans. Pandas

* 1. **Data Acquisition:** 
     1. **Open Source data (or) via API (or) via scrapping (or) Generate synthetic Data**

Ans. Open source data

* + 1. **Can you acquire more data in addition to your primary source? If so, how?**

Ans. I think more should be available to us by searching Google.

1. **Key metric (KPI) to optimize**
   1. **Business Metric definition.**

Ans. If a chatbot returns “I don’t know (IDK)” for the questions about which it is not aware. Users have the option to request a callback by human agent in case they are not satisfied with the chatbot. At the end of the session, the chatbot requests users for a rating on a scale of 1-5.

Then, Performance of the chatbot can be measured using the following metrics-

1. Callback ratio = No. of users requested callback from human agent per day / Total no. of users per day
2. IDK ratio = No. of questions for which chatbot responded as IDK per day / Total no. of questions asked per day
3. Mean of all the ratings in a day = Sum of all the ratings / Total no. of rating received in a day
   1. **Why is this metric used ?**

Ans. 1) Callback ratio can be used to measure the effectiveness of a chatbot as lower the callback ratio, the more satisfied the customers are. 0 <= Callback ratio =< 1. The higher the callback ratio, the more dissatisfied the customers are.

2) IDK ratio can be used to measure whether the users are getting the answers of their questions or not. 0 <= IDK ratio =< 1. Lower IDK ratio means less no. of questions went unanswered. The higher the IDK ratio, the more are the questions that went unanswered.

3) Mean of all the ratings- Overall user experience can be known using the mean of all the ratings. If the mean is good (e.g. >3), then users are having a good time with the chatbot. If the mean is poor ,like, <=2, then it means users are dissatisfied with the chatbot.

* 1. **Alternative metrics that can be used? Why are they not preferred in this case?**

Ans. How many individuals are needed to answer the number of questions that the chatbot did- Let us suppose that it takes on an average 3 minutes to handle one question by a human agent. And a human agent spends 5 hours on a daily basis to answer user queries. So, a human agent can answer 100 questions daily. Let 1000 questions are being asked on a daily basis. Then, it takes 10 human agents to answer all these questions. Hence, using chatbot saved the cost of 10 human agents.

This metric is not preferred in this case because this doesn't measure customer satisfaction and areas of improvement for chatbot.

* 1. **Pros and cons of the metric used.**

Ans. 1) Callback ratio:

Pros- Callback ratio can be used to measure the effectiveness of a chatbot as lower the callback ratio, the more satisfied the customers are. 0 <= Callback ratio =< 1. The higher the callback ratio, the more dissatisfied the customers are.

Cons- This is not necessary that users are requesting callbacks due to the fact that the chatbot was unable to satisfy them. Users might be more comfortable interacting with the human agent, so they are requesting callbacks without even trying a chatbot. So, by using only callback ratio, we cannot say with certainty that these are the no. of users who got dissatisfied with the chatbot.

2) IDK ratio:

Pros- IDK ratio can be used to measure whether the users are getting the answers of their questions or not. 0 <= IDK ratio =< 1. Lower IDK ratio means less no. of questions went unanswered. The higher the IDK ratio, the more are the questions that went unanswered.

Cons- Chatbot can respond with IDK in two cases: First, in which it is not aware of the answer for the valid question which is asked by the user. Second, in case the user intentionally asks invalid questions (which chatbot is never supposed to know, e.g. Where does the sun rise from?). We have to remove the no. of second cases in order to get the true value for the no. of first cases.

3) Mean of all the ratings in a day:

Pros- Overall user experience can be known using the mean of all the ratings. If the mean is good (e.g. >3), then users are having a good time with the chatbot. If the mean is poor ,like, <=2, then it means users are dissatisfied with the chatbot.

Cons- For users who are not interested in using the chatbot, and are just using the chatbots to get the callbacks. For these users, the ratings given by them do not reflect the true user experience of the chatbot.

* 1. **Where does this metric fail? Where should it not be used?**

Ans. Callback ratio- This is not necessary that users are requesting callbacks due to the fact that the chatbot was unable to satisfy them. Users might be more comfortable interacting with the human agent, so they are requesting callbacks without even trying a chatbot. So, by using only callback ratio, we cannot say with certainty that these are the no. of users who got dissatisfied with the chatbot.

IDK ratio- Chatbot can respond with IDK in two cases: First, in which it is not aware of the answer for the valid question which is asked by the user. Second, in case the user intentionally asks invalid questions (which chatbot is never supposed to know, e.g. Where does the sun rise from?). We have to remove the no. of second cases in order to get the true value for the no. of first cases.

Mean of all the ratings in a day- For users who are not interested in using the chatbot, and are just using the chatbots to get the callbacks. For these users, the ratings given by them do not reflect the true user experience. So, these ratings are useless.

* 1. **Where (on what type of problems is this metric used elsewhere in ML and Data Science?**

Ans. Mean of all the ratings can be used in almost any type of problems. Callback ratio and IDK ratio can be used in problems related to the chatbot.

* 1. **Code: Implement this metric from scratch in Python using** **NumPy/Pandas/Scipy only.**

1. **Real world challenges and constraints**
   1. **What real world constraints do you have while solving this problem?**

Ans. 1) Users are comfortable to talk to a human representative over chatbot. Due to this reason, extra efforts need to be put into making the user try chatbot first, and if he/she does not find the satisfying answer using chatbot, then turn to human representatives.

* 1. **What are the requirements that your solution must meet?**

Ans. 1) Must provide the relevant answer for the question asked by the user.

2) Must provide a quick response to the user.

3) If a question asked by the user is vague, then the chatbot must return the more concise form of the question asked by the user and an answer corresponding to it. Chatbot must also confirm from the user if he meant the same question as responded by the chatbot.

4) Must return the response as "Sorry, I don't understand you." for the question for which chatbot does not have an answer.

5) Count the no. of times the chatbot has returned "Sorry, I don't understand you", so that these numbers can be used to measure the efficiency and improve the knowledge of the chatbot.

6) Count the no. of times the user has chosen to escalate to the human representative, so that these numbers can be used to measure the efficiency and improve the knowledge of the chatbot.

7) At the end of the session, the chatbot must request users for a rating on a scale of 1-5.

1. **How are similar problems solved in literature ( technical articles & blogs, forums, research papers)** 
   1. **How is the problem mapped to an existing ML methodologies: Classification/ Regression/Time-series-forecasting/DeepLearning/RecSys etc?**

Ans. 1) A state of the art dialog agent has been built by using Transfer Learning based on OpenAI GPT and GPT-2 Transformer language models.

2) Chatbots addressing mental health are built on the premise of cognitive-behavioural therapy (CBT). CBT uses structured exercises to encourage a person to question and change their habits of thought— this format is well suited to a step-by-step software guide or chatbot.

3) Woebot, founded in 2017 by a team of Stanford psychologists and AI experts, is a talk therapy chatbot helping its users monitor their mood and learn about themselves. They use Artificial Intelligence (AI) and Natural Language Processing (NLP) techniques to understand and adapt to a person’s emotional state and cognitive ability, to motivate them to monitor their mood, and to practice proven therapeutic techniques.

4) Karim is an example of an Arabic-speaking chatbot, designed by the Silicon Valley startup X2AI that helps Syrian refugees deal with their mental health issues.

* 1. **List all solution approaches that you think are relevant to your problem.**

Ans. 1) Transfer learning combined with OpenAI GPT and GPT-2 Transformer language models.

2) Rasa can be used.

* 1. **List/Cite all references you go through.**

Ans.

1) https://www.thoughtworks.com/en-in/insights/blog/how-build-excellent-chatbot

2) https://medium.com/analytics-vidhya/building-a-simple-chatbot-in-python-using-nltk-7c8c8215ac6e

3) https://medium.com/huggingface/how-to-build-a-state-of-the-art-conversational-ai-with-transfer-learning-2d818ac26313

4) https://www.lyntonweb.com/inbound-marketing-blog/16-kpis-to-measure-chatbot-effectiveness

5) https://woebothealth.com/

6) https://digital.hbs.edu/platform-digit/submission/woebot-the-bleeding-intelligent-self-help-therapist-and-companion/

7) https://analyticsindiamag.com/chatbots-in-mental-health-friendly-but-not-too-friendly/

8) <https://www.analyticsvidhya.com/blog/2018/01/faq-chatbots-the-future-of-information-searching/>

9) https://link.springer.com/chapter/10.1007/978-3-030-47426-3\_19