WGU C951

Task 3

MACHINE LEARNING PROJECT PROPOSAL

Jacob D. Sanchez

Student ID # 001242735

October 28, 2023

**A. Project Overview**

This new machine learning system is designed to aid technical support agents in swiftly addressing customer inquiries and identifying emerging trends.

**A.1. Organizational Need**

Our organization delivers top-tier technical support for WordPress websites, and we are currently facing challenges in managing the increasing customer demand. This growth in customer contacts is leading to longer wait times in our chat queues. With an internal AI chatbot, I intend to reduce overall chat queues and reduce the stress of fellow support agents.

**A.2. Context and Background**

Our organization has been around for more than 10 plus years and have always strived to deliver quality support. With the help of an internal AI chatbot, we can quickly get answers to questions that are accurate and reduce overall chat queues.

**A.3. Outside Works Review**

One example of an organization that did something similar is found in an article called “How a leading telco increases client & employee satisfaction with conversational AI” (n.d) states “the company saved an average of 10-15 minutes per customer support interaction – a drastic improvement with an undeniably high ROI”. Utilizing a chatbot is essential for growing business with high customer contacts that also has high reward.

Another example of an organization comes from article “How iFood served up conversational AI to improve B2B communication with drivers, restaurants, and end users” (n.d), stated “The AI solution handles and automates 45% of incoming inquiries resulting in 70% in reduction in delivery costs”. They were able automate tasks coming from customer inquiries and reduce in overall delivery costs thanks to the AI chatbot.

Third example of an organization comes from article “How Argenta banked a 95% CSAT score with an AI chatbot” (n.d), states “Charlie handled 20% of incoming conversations, which said 24 hours of work for agents in the first one and a half months alone! It’s important to note that if Charlie doesn’t know an answer, the conversation gets routed to a live agent, so the same high-level quality conversation is support”. Here we can see how AI chatbots can be beneficial in increasing or maintaining CSAT and reducing overall contacts that had to be handled to support agents.

**A.4. Solution Summary**

With this new AI chatbot, our organization will be able to reduce chat queues and answer technical questions much faster. This will help our support agents not feel as much burden and provide high quality support with lower volume of customer contacts. For us to accomplish this task, we will primarily focus on how to integrate an internal chatbot system with existing software.

**A.5. Machine Learning Benefits**

Given the technical advancements in AI in the past few years, it’s becoming apparent that AI can answer and provide solutions faster than people can. With this AI, it will be able to analyze technical inquiries and provide information related to the topic much more quickly that may not be obvious to a support agent.

**B. Machine Learning Project Design**

**B.1. Scope**

**In Scope**

* This project will see the installation of AI chatbot integrated with existing software.
* The in-house developed AI will gather certain keywords used and provide technical answers in a conversational matter.
* This new system will be able to track trends in topics from support agents and continue to learn to provide valuable information regarding the issues customers are experiencing.

**Out of Scope**

* This new system will not be available for customer use and will not keep track of who asked questions within the AI chatbot.

**B.2. Goals, Objectives, and Deliverables**

Goals

• Create an AI chatbot that can provide valuable information to support agents in hopes to provide faster response times and lower chat queue volume.

Objectives

• The algorithm will be able to identify certain keywords or patterns to provide valuable information to support agents.

Deliverables

• An AI chatbot developed in-house that can help support agents answer customer technical inquires much faster based on certain keywords or patterns.

**B.3. Standard Methodology**

Development will follow the SEMMA methodology.

• Sample: The data we will gather for this chatbot AI system will be in two parts, feeding the AI data from current knowledge cards and plan on gathering 1000 chat examples to review top topics.

• Explore: We will review the data in-house and determine any top topics or patterns for the AI as well as if there are any conflicts in the conversation that may cause issues for the AI to learn. For example, gibberish, blank text, or special characters.

• Modify: Once we have been able to explore the data, we will now go into the modify phase. All gibberish, special characters or blank input will be ignored. We will also write additional code that will try to “guess” what the user is trying to say to train the AI further in less ideal circumstances.

• Model: We will then give this data to the AI to continue learning new topics or patterns to provide relevant information to the user with the initial inquiry for best accuracy and results

• Assess: Once the AI has setup for additional training, we will then test for accuracy and efficiency to determine top topics and if the AI is providing relevant information. We will do this by selecting 100 support agents and ask the AI questions to see if the AI provides accurate results as customer contacts come along.

**B.4. Projected Timeline**

November 1st, 2023 – Proposal is accepted.

November 2nd, 2023 – A technical proof of concept is presented.

November 3rd, 2023 – Submitted for review…

January 1st, 2024 – Deliverables

January 15th, 2024 – Delivered.

**Sprint Schedule**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sprint** | **Start** | **End** | **Tasks** |
| 1 | November 1st, 2023 | November 15th, 2023 | 1. Define the project scope and objectives. 2. Assemble engineering and project team. 3. Configure hardware and software. 4. Data collection from knowledge cards and clean the data. 5. Develop the AI chatbot with the NER algorithm using Python and PyTorch 6. Create UI for chatbot |
| 2 | November 16th, 2023 | November 30th, 2023 | 1. Continue data collection and cleaning. 2. Utilize NER algorithm to extract and categorize entities from knowledge cards. 3. Train chatbot on a set of data 4. Identify keywords and patterns. 5. Start testing the AI chatbot with a group of agents. 6. Review and update data based on feedback. |
| 3 | December 1st, 2023 | December 15th, 2023 | 1. Chatbot continues to be trained on more data. 2. Improve the chatbots ability to detect keywords and patterns. 3. Chatbot to identify gibberish or special characters. 4. Continue testing with support agents. 5. Identify any limitations during testing. |
| 4 | December 16th 2023 | January 1st, 2023 | 1. Finish up AI chatbots development. 2. Further testing of accuracy and efficiency 3. Identify if the objectives were met based on the results. 4. Finalize the project deliverables. 5. Train agents how to use the new AI chatbot system. 6. Make any final adjustments before the release. |
| 5 | January 15th, 2023 | January 15th, 2023 | AI chatbot system is delivered. |

**B.5. Resources and Costs**

|  |  |  |
| --- | --- | --- |
| **Resource** | **Description** | **Cost** |
| Software Engineers | To develop in-house AI chatbot system and implementing required algorithm | $100,000 |
| Data Engineers | To review and modify existing data and works with software engineers for integration | $150,000 |
| Server (Hardware) | This will host the AI chatbot | $10,000 |
| Ubuntu OS (Software) | The AI chatbot will run on server hardware running Ubuntu | $0 |
| Data Collection | Data is provided in-house (internal knowledge cards) | $0 |
| PyTorch Tooling | Open-source software to build and train the AI | $0 |
|  | **Total** | $260,000 |

**B.6. Evaluation Criteria**

Describe the criteria used to evaluate and measure the success of the completed project.

|  |  |
| --- | --- |
| **Objective** | **Success Criteria** |
| Algorithm accuracy | The algorithm will be tested with 100 agents to test different conversational questions to the chatbot AI system that will be used to train it. If the algorithm has a success rate of 85% from the samples, then it will pass. |
| Data Accuracy | The new system will be successful if it can report accurate data to support agents from November 3rd to January 1st 2024 |
| Ease of use | The AI chatbot system will be successful if support agents are able to start chatting within two minutes. |

**C. Machine Learning Solution Design**

**C.1. Hypothesis**

Our organization is having problems keeping up with the demanding chat queues and is becoming an overwhelming task for our support agents. I believe that with this new internal AI chatbot, we can answer technical questions much faster and reduce overall chat queues. It would also allow us to maintain a high CSAT.

**C.2. Selected Algorithm**

For this AI chatbot system, we will be using Named Entity Recognition algorithm (NER). An NER is an algorithm that involves identifying and categorizing entities within text. This would allow us to extract certain pieces of text and provide information based on that entity.

**C.2.a Algorithm Justification**

I chose this algorithm due to its popularity and functionality. According to Chacko (2023), “NER is pivotal in supercharging customer care functions. It helps an AI tool automatically categorize queries and complaints by identifying keywords (such as brand names or branch locations), so they’re queued and routed to relevant customer care teams for smoother support”. Here we can see how beneficial NER is for AI chatbots.

**C.2.a.i. Algorithm Advantage**

There are many benefits of NER algorithm. One reason is that they enhance information retrieval and search accuracy. This is excellent because it can identify and categorize entities in text which will provide more precise information back to the user.

**C.2.a. ii. Algorithm Limitation**

One limitation of the NER algorithm is that when there are many entities or variations in the text, this could through off the results and requires additional training.

**C.3. Tools and Environment**

This new AI chatbot system will be developed and trained with Python, a programming language. We will be using PyTorch as this is an excellent open-source tool that supports the development and training of AI.

**C.4. Performance Measurement**

The quality and performance of the algorithm will be measured by the accuracy of the answers given to the support agents depending on certain keywords or patterns. We will do this by providing the AI a collection of knowledge cards and dedicating 100 support agents to test different questions and topics which then a support agent can provide feedback of the results. We will then examine the data/results to determine its accuracy.

**D. Description of Data Sets**

**D.1. Data Source**

The data sources will be from the internal organization. The first piece of data will be from internal knowledge cards that support agents use daily. These cards contain various topics on certain customer facing issues. We will first have to train the AI of this set of data. The second piece of data is collecting top topics from support agents to better gauge customer facing issues.

**D.2. Data Collection Method**

The organizations data collection method is using existing internal knowledge cards and feeding that to the AI chatbot system. Topics will also be tracked internally with the help of the AI. No outside resources will be required in this case.

**D.2.a.i. Data Collection Method Advantage**

The advantages of using this data collection method are that lot of issues have been recorded in an internal knowledge card. This saves a lot of time and resources having to build a new data set that the AI must pull from. The second advantage is we can identify any trends of topics asked by support agents.

**D.2.a. ii. Data Collection Method Limitation**

The limitation of this data collection method is if there is a new topic or issue that has not been recorded in a knowledge card, may yield inaccurate results.

**D.3. Quality and Completeness of Data**

The quality and completeness of the data should be sufficient as these are regularly updated and maintained. However, we will prepare this data by removing any redundant information as well as for any grammatical errors that may cause issues for the AI.

**D.4. Precautions for Sensitive Data**

The data in the current knowledge cards are not restricted as its open to all of support. However, we will take precautions to ensure support agents are not releasing any proprietary information to customers by having them review answers given from the AI.

**References**

Chacko, A. (2023, November 1). *How named entity recognition (NER) helps marketers discover Brand insights*. Sprout Social. <https://sproutsocial.com/insights/named-entity-recognition/>

*How a leading telco increases Client & Employee Satisfaction With Conversational AI*. Sinch. (n.d.). <https://www.sinch.com/insights/customer-stories/telco-increases-client-employee-satisfaction-conversational-ai/>

*How ifood served up conversational AI to improve B2B communication with drivers, restaurants, and end-users*. Sinch. (n.d.-b). <https://www.sinch.com/insights/customer-stories/how-ifood-served-conversational-ai-improve-b2b-communication-drivers/>

*How Argenta banked a 95% CSAT score with an AI chatbot*. Sinch. (n.d.-b). <https://www.sinch.com/insights/customer-stories/how-argenta-banked-95-csat-score-ai-chatbot/>