**Text Analytics Final Project**

**Team Members:**

**Ganti Uday**

**2**

**3**

**4**

**5**

**Executive Summary**

The purpose of this report is to provide a detailed account of our initial meeting with Dr. Han Reichgelt, held on September 27, 2023. The meeting, which lasted for approximately 20 minutes, was focused on discussing the prospects of implementing a chatbot using an LLM to create a chatbot designed to handle student queries efficiently. Over this period, we discussed the current pipeline, the options for automation that currently exist, the pain points and redundancies, and what tasks would improve efficiency or reduce latency if automated.

**Business Process Analysis**:

Detail the existing MSBAIS business process landscape, identify bottlenecks, and explore areas where a ChatGPT-like interface could add value.

A diagram of a diagram

Description automatically generated

Current BAIS Student Query Workflow Diagram

**Workflow:**

The MS-BAIS program at the University of South Florida has a specific workflow for handling incoming queries via their email address "ms-bais@usf.edu".

Initially, all emails received are converted into "tickets" using Gira Software, which is likely a misspelling of Jira, a commonly used ticketing software. These tickets are then categorized based on 'keywords' to determine whether they are of an administrative or academic nature.

Once categorized, each email undergoes a manual evaluation to assess its content and decide the next steps.

Depending on the evaluation, emails are either redirected to the relevant personnel for further action, replied to directly, or, in cases where the query is not clear, a request for additional information is sent back to the sender.

**Pain Points**

* While the process is structured, it has some potential drawbacks. For instance, manual evaluation can be time-consuming, especially if crucial information such as U ID isn’t mentioned.
* Most frequent queries are solved by redirecting students to the “Canvas Current Student Portal”.

**Client Needs and Objectives:**

While there will always be a need for manual intervention for specialized queries and administrative decisions, the process could be streamlined by making sure there is an automated method of making sure all the necessary information is made available.

Additionally, repeated and frequent queries with simple default responses can have automated replies.

Another area of interest for automation is the queries, which are solved by redirecting the student to the relevant “Canvas Current Student Portal”.

**Data Assessment**:

The data made available for this project is 5,000 emails sent through the ticketing system.

Given that emails can contain sensitive information, considerations about data privacy and confidentiality must be addressed; hence, the data provided is anonymous.

Data Limitations:

* Volume: While 5,000 emails provide a sizable dataset, it may not be enough to capture the full range of student queries.
* Variability: The dataset may have inherent biases based on the time period it covers; for example, more queries about admissions may be received during a specific season.

**Scope of Use**:

* Automated Information Gathering: A chatbot can scan the email content for missing crucial information like the University ID (U ID) and automatically request it from the sender.
* For frequent and straightforward queries, the chatbot can be trained to automatically provide relevant information, such as redirecting students to the Canvas Current Student Portal.
* Machine learning algorithms can be trained to improve upon the existing keyword-based categorization, making it more accurate and efficient.

**Next Steps & Follow-up activities and action items**

Moving forward, the immediate next step is to conduct a thorough data cleaning and preprocessing operation to make the 5,000 emails suitable for analysis and model training.

Once the data is preprocessed, a pilot phase should be initiated to train initial machine learning models for categorizing queries and automating responses.

We should also schedule a follow-up meeting with Dr. Han Reichgelt to present these initial findings and plans and to discuss any adjustments to the project scope or objectives.

Furthermore, a plan should be laid out for continuous monitoring and improvement, which will include metrics to evaluate the chatbot's effectiveness in streamlining the query management process.

**Conclusion**

In conclusion, the initial meeting with Dr. Han Reichgelt and the subsequent analysis laid a strong foundation for the implementation of a chatbot system aimed at streamlining the MS-BAIS program's query management process at the University of South Florida. The availability of a sizable dataset of 5,000 emails, along with a detailed understanding of the current workflow and its pain points, provides us with a unique opportunity to create a meaningful impact.

Through careful data assessment, targeted machine learning algorithms, and strategic integration into existing systems, the project aims to reduce manual workload, increase efficiency, and enhance the overall experience for both students and administrative staff. By adhering to the outlined next steps and follow-up activities, the project is well-positioned to meet its objectives in a timely and effective manner.