NLP Course: CRM summarizer

Fedor Dobryansky, Bohdan Bulatov, Dmitriy Didenko

December 2023

Abstract

"Crmsum" is a tool designed to improve sales strategies by aggregating and summarizing sales interactions from a CRM system. It identifies patterns of success and failure, providing sales teams with actionable insights to enhance performance. `https://github.com/mlenzovet/crmsum`.

1 Introduction

Understanding the factors that lead to successful sales is a critical yet challenging task for many businesses. The necessary data is often scattered and unstructured within Customer Relationship Management (CRM) systems. "Crmsum" addresses this problem by providing an automated solution to extract, analyze, and summarize sales interactions from a CRM system.

What sets "Crmsum" apart is its simplicity and effectiveness. It uses Natural Language Processing (NLP) to parse user queries, accesses the CRM via an API, and provides a clear summary of a salesperson's interactions over a specified period. This approach enables sales teams to learn from past experiences and enhance future performance in a practical and efficient way.

1.1 Team

- **Fedor Dobryansky**
- **Bohdan Bulatov**
- **Dmitriy Didenko**

2 Related Work

Prior research in the field of Customer Relationship Management (CRM) predominantly centers around customer satisfaction assessment through sentiment analysis. Lexicon-based and corpus-based approaches have been widely explored [1]. Lexicon-based methods utilize dictionaries of positive and negative words to estimate sentiment, considering local context factors. Corpus-based approaches, on the other hand, build sentiment classifiers from annotated sentences, often incorporating additional information like emoticons or star ratings.

While social media data has been extensively analyzed for its impact on organizations, there remains a notable gap in the literature regarding sentiment analysis applied to direct communications within CRM systems. Existing studies have employed various techniques, including lexicon-based approaches and clustering, to analyze emails and other forms of direct communication. However, few have focused on incremental learning capabilities, which allow models to adapt to new concepts without retraining the entire classifier.

Another study [2] introduces an approach utilizing Hierarchical Attention Networks (HAN) for sentiment analysis of customer communications within Customer Relationship Management (CRM) systems. Unlike existing methods, this model incorporates an integrated incremental learning mechanism, allowing it to improve over time with feedback from CRM operators. The research demonstrates the effectiveness of this approach with a dataset of over 30,000 annotated items, achieving superior sentiment classification performance compared to other methods. The proposed methodology addresses the challenge of efficiently processing and prioritizing customer communication stimuli in modern CRM systems, emphasizing the importance of automated language processing techniques.

In one more example [3] the study focuses on utilizing a text corpus stored in a Customer Relationship Management (CRM) database for data mining and segmentation. It leverages established methods from Natural Language Processing (NLP) and deep learning, including word embeddings and recurrent neural networks (RNNs), to analyze text notes from a CRM system used by customer representatives of an internet ads consultancy agency over a decade. The research demonstrates that word embeddings derived from the text corpus can be directly applied for data mining purposes and integrated into RNN architectures with long short-term memory (LSTM) units for more extensive segmentation tasks. The results affirm the feasibility of extracting valuable insights from structured text data within a CRM system. This underscores the potential for implementing NLP features in any CRM platform, provided that problem definitions are accurately formulated and solution methods are appropriately implemented.

In contrast, this study takes a distinctive approach by concentrating on the NLP analysis of CRM activity. By applying advanced NLP techniques to CRM interactions, this research aims to uncover valuable insights into project performance and manager productivity. This innovative approach offers the potential to provide data-driven perspectives that can enhance decision-making and optimize CRM strategies for improved project outcomes.

Since we're implementing LangChain in out project, it is reasonable to utilize the following research [4] which introduces a groundbreaking approach to automating customer service using LangChain, a custom Large Language Model tailored for organizations. The study highlights the obsolescence of traditional customer support techniques, particularly Frequently Asked Questions (FAQs), and advocates for a shift towards more responsive, context-aware, and personalized customer interactions. The integration of LangChain into customer service platforms, presented as the open-source framework "Sahaay," demonstrates its ability to scale across industries and organizations, providing real-time support and query resolution. The research also emphasizes the role of web scraping, embeddings, and the utilization of Google's Flan T5 XXL, Base, and Small language models for knowledge retrieval. The results section offers insights into performance and use cases, particularly within an educational institution. This research marks a new era in customer service, leveraging technology to create efficient, personalized, and responsive interactions, ultimately redefining the customer-company relationship for enhanced customer retention, value extraction, and brand image.

In the paper titled "Revisiting Prompt Engineering via Declarative Crowdsourcing," [5] the authors propose a more principled approach to prompt engineering for large language models (LLMs) by drawing inspiration from declarative crowdsourcing techniques. They treat LLMs as analogous to crowd workers and advocate for strategies such as employing multiple prompts, ensuring internal consistency, and exploring hybrid LLM-non-LLM approaches. The first proposed approach involves varying prompting strategies by breaking down complex questions into smaller tasks, allowing for a more detailed and nuanced response. The second approach suggests a hybrid coarse to fine-grained prompting, where tasks are sequentially completed at different levels of granularity to obtain the final result. Lastly, the authors advocate for leveraging both LLM and non-LLM approaches, emphasizing cost reduction by using cheaper alternatives for certain tasks instead of relying solely on resource-intensive language models. This may include employing vector search on knowledge bases or utilizing methods like nearest neighbors when the full capacity of LLMs is unnecessary for certain queries.

3 Model Description

The model developed, named *crmsum*, operates in several steps to extract, analyze, and summarize sales experience from the AMO CRM system using the capabilities of the ChatGPT API.

1. **Query Processing:** The model first processes the user's query submitted via a Telegram chat. The query is expected to contain the name of a salesperson (manager), the name of a company, and a date range. The model uses the ChatGPT API to extract these pieces of information from the query. Specifically, the model uses functions such as `find_manager_name`, `find_customer_name`, and `find_date_range` to perform this extraction.

2. **Record Retrieval:** Once the manager's name, company name, and date range have been extracted from the user's query, the model then communicates with the AMO CRM system's API to retrieve relevant records. This is done through the `get_records_data` function, which accepts the extracted information as input along with a token and an API path. This

function returns a list of texts from the records that match the specified filters (manager name, company name, date range).

3. **Record Summarization:** After the relevant records have been retrieved, the model uses the ChatGPT API again to summarize the texts from these records. The `summarize_records_texts` function accepts the list of texts as input and returns a list of summarized texts.

Overall, the model provides a streamlined way to summarize the sales experience of a specific manager with a particular company over a defined period of time, offering valuable insights for improving sales performance.

4 Dataset

The primary datasets used for the development and evaluation of the *crmsum* model are sourced from two different platforms: AMO CRM system and a Telegram chat history.

1. **AMO CRM Dataset:** This dataset consists of customer request cards from the AMO CRM system. Each card contains between 10 and 200 records, amounting to a total of approximately 1000 cards. Each record consists of textual data and metadata related to customer interactions, sales activities, and outcomes.

2. **Telegram Chat History Dataset:** This dataset comprises a 3-year history of a Telegram chat with salespeople. The chat contains between 5 and 100 messages per day, providing a rich source of informal, conversational text and sales communication data.

These datasets were collected and utilized in accordance with appropriate data protection and privacy regulations. While the specifics of the data make it unsuitable for public release, the methods and techniques applied in this study can be replicated on similar datasets.

	AMO CRM Dataset	Telegram Chat Dataset
Source	AMO CRM	Telegram
Number of Cards/Messages	1000	3 years of daily messages
Range of Records per Card/Messages per Day	10-200	5-100

Table 1: Overview of the datasets used in the study.

The collection procedure for the datasets involved the extraction of data via the respective platform APIs. For the AMO CRM dataset, this involved the retrieval of customer request cards and their associated records. For the Telegram dataset, this involved the extraction of chat history over a specified period. In both cases, the data was pre-processed and anonymized to ensure privacy and compliance with data protection regulations.

The AMO CRM data provides a detailed account of sales activities and outcomes, offering valuable context for understanding and improving sales performance. The Telegram chat history offers a conversational perspective on sales communications, providing insights into the informal and unstructured interactions that characterize sales discussions.

5 Experiments

5.1 Metrics

The primary metric for evaluating the success of the system will be feedback from the salespeople themselves, collected through a rating system in the chat interface. Salespeople will be asked to rate the quality and usefulness of the summaries on a scale from 1 to 5.

A secondary, and more long-term, measure of success will be an observed increase in sales volume over the course of a year. This will provide a quantitative measure of whether the system is improving the effectiveness of the sales team.

There also metrics for the RAG evaluation [8] to be utilized:

- **Faithfulness**: This metric evaluates how accurately language model responses reflect information from the data sources, ensuring that the model does not distort or alter information during presentation.
- **Answer Relevancy**: It indicates the extent to which model responses align with the posed questions. This is crucial, as even technically correct answers may be useless if they do not address the specific user query.
- **Context Recall and Context Precision**: These metrics assess how well the model utilizes question context in generating responses. "Context Recall" evaluates how much information from the context is incorporated into the answer, while "Context Precision" gauges the accuracy and relevance of contextual information used.
- **Answer Semantic Similarity**: This metric measures how closely the model's response aligns semantically with a reference answer, providing insights into how naturally and accurately the model can reproduce human language.
- **Answer Correctness**: It evaluates the accuracy of the model's response, serving as the final check to determine if the model provides precise and reliable information in its answers.

More detailed overview and implementation could be checked at the gihub page[9].

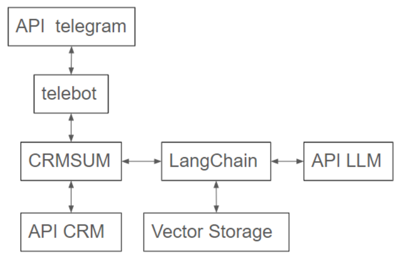

Figure 1: CRMSUM modules

5.2 Experiment Setup

The experiment will be conducted over the course of a year. The system will regularly generate summaries of sales experiences, which will be delivered to the sales team through the chat interface. Salespeople will be encouraged to provide feedback on each summary through a rating system. The ratings will be collected and analyzed to evaluate the performance of the system over time.

In addition, the total sales volume will be tracked over the course of the year. An increase in sales volume compared to the previous year will be taken as an indication that the system is improving the effectiveness of the sales team.

5.3 Baselines

As this is a novel application with no previous art, there are no established baselines for comparison. However, the performance of the sales team without the system, as measured by sales volume over the past year, will be used as a baseline for comparison.

5.4 Modular integration

As shown in [Figure 1]

- **API Telegram**: Interaction: Bidirectional communication with 'telebot.' Purpose: Facilitates communication with the Telegram messaging platform, allowing for the sending and receiving of messages.

- **Telebot**:

  Interaction: Bidirectional communication with 'API Telegram' and 'CRMSUM.' Purpose: Serves as a bridge between the Telegram messaging platform and the system components, enabling communication with both users via Telegram and internal components.

- **CRMSUM**:

Interaction: Bidirectional communication with 'telebot' and 'API CRM.' Purpose: Manages and summarizes CRM data, providing relevant information to users through 'telebot' and interacting with the CRM system via 'API CRM.'

- **API CRM**:

  Interaction: Bidirectional communication with 'CRMSUM.' Purpose: Connects the system with the CRM (Customer Relationship Management) system, allowing the retrieval and updating of CRM data.

- **Langchain**:

  Interaction: Bidirectional communication with 'CRMSUM,' 'API LLM,' and 'Vector Storage.' Purpose: Handles natural language processing tasks, interacting with 'CRMSUM' for contextual understanding, 'API LLM' for language model operations, and 'Vector Storage' for efficient data storage.

- **API LLM** (Language Model):

  Interaction: Bidirectional communication with 'Langchain.' Purpose: Provides access to language models for natural language understanding and generation, enhancing the system's ability to comprehend and respond to user inputs.

- **Vector Storage**:

  Interaction: Bidirectional communication with 'Langchain.' Purpose: Stores and retrieves vector representations efficiently, supporting 'Langchain' in processing and analyzing data in a vectorized format. This integration allows for a robust and interconnected system, where each component plays a specific role in ensuring effective communication, data processing, and natural language understanding within the overall system architecture.

5.5 Utilizing Best Practice

For a successful integration and efficient CRM data analysis using GPT and LLM, we adhere to the following best practices [6, 7]:

1. **Data Privacy and Security Measures**: In the integration of GPT and LLM for CRM data analysis, a paramount focus has been placed on data privacy and security. This involves the secure handling of customer data and strict adherence to relevant data protection regulations, such as GDPR or HIPAA, where applicable.

2. **Ethical AI Usage Oversight**: To maintain ethical AI usage, there have been regular reviews and fine-tuning of GPT and LLM responses. This ensures that the models do not generate inappropriate or biased content, with clear guidelines established to guide their behavior during CRM data analysis.

3. **Training and Support Initiatives**: Meetings and support initiatives have been implemented for team members engaged in interacting with GPT and LLM for CRM data analysis. This includes encouraging to share feedback on the experiences and promptly addressing any challenges encountered during the process.

4. **Implement a Feedback Mechanism**: We plan to set up a feedback mechanism to continuously enhance the system. Gather insights from both users and administrators to make informed, data-driven decisions for ongoing optimization.

5. **Ensure Scalability and Performance**: Measures have been taken to ensure the scalability and performance of the integration. This involves monitoring the system's capability to handle increased demand as the organization grows, with proactive steps taken to prevent bottlenecks during periods of peak usage.

6 Conclusion

In this work, we have described the design and implementation of a novel system, crmsum, for summarizing and sharing sales experiences. The system aims to leverage the wealth of information contained in CRM records to improve the effectiveness of sales teams. By providing concise and relevant summaries of past sales experiences, the system enables salespeople to learn from each other's successes and failures, and apply these lessons to their own sales efforts. We have outlined the system's functionality, described the algorithms used for extracting and summarizing information, and proposed an experimental design for evaluating the system's performance. While the experiment has not yet been conducted, we believe that crmsum has the potential to significantly enhance sales performance and drive business growth.

References

The following links are provided:

1. [https://wires.onlinelibrary.wiley.com/doi/10.1002/widm.1171]

2. Nicola Capuano, Luca Greco, Pierluigi Ritrovato, Mario Vento, "Sentiment analysis for customer relationship management: an incremental learning approach" 2020

3. Şükrü Ozan, "Case studies on using natural language processing techniques in customer relationship management software", 2023

4. Keivalya Pandya, Prof. Dr. Mehfuza Holia, "Automating Customer Service using LangChain", 2023

5. Aditya G. Parameswaran, Shreya Shankar, Parth Asawa, Naman Jain, Yujie Wang, "Revisiting Prompt Engineering via Declarative Crowdsourcing", 2023

6. Olga Green, "Revolutionizing CRM with GPT and LLM: A Comprehensive Overview"

7. Monu Kumar, "How GPT and LLM are Revolutionizing CRM Software?"

8. Оцениваем RAG-пайплайны

9. RAGAS - Evaluation framework for your Retrieval Augmented Generation (RAG) pipelines