International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V72I7P110

Volume 72 Issue 7, 83-86, July 2024

Original Article

Boost Call Center Operations: Google's Speech-to-Text AI Integration

Suman Chintala

Business Intelligence Architect, 66degrees LLC.

Corresponding Author : chintalasuman@outlook.com

Received: 16 May 2024 Revised: 27 June 2024 Accepted: 17 July 2024 Published: 31 July 2024

Abstract - *A positive customer service experience can significantly boost a company's growth by fostering customer loyalty, reducing customer churn, and strengthening brand value. However, improving the efficiency of this process through automation and AI/ML still needs to be explored. This paper presents a solution for call centers that enhances customer support efficiency and uncovers business insights from audio data by integrating Google's Speech-to-Text API with the call center database (Big Query) and conducting post-call analytics using Looker Studio, aiming to revolutionize call center operations by enhancing efficiency and uncovering valuable business insights.*

Keywords - *Artificial Intelligence, Generative AI, Machine Learning, Business intelligence, Call center operations, Speech-to-Text, Sentiment analysis, Google cloud, Looker studio, Data analytics.*

1. Introduction

A positive customer service experience can significantly boost a company's growth by fostering customer loyalty, reducing customer churn, and strengthening brand value. Hence, companies today focus on improving their customers' customer service experience. However, more focus has been needed to improve the efficiency of this process through automation and deploying AI/ML to generate meaningful customer insights, uncover new business opportunities, and improve call center efficiencies.

This article presents a solution for call centers that not only boosts the efficiency of customer support executives but also holds the promise of uncovering business trends and insights from audio data. This is accomplished by integrating Google's Speech-to-Text API with the call center database (Big Query) and conducting post-call analytics using Looker Studio.

2. Background/Context

2.1. Current Process

The current process at a call center is such that the customer care executive receives the call from the customer. These calls are recorded in an audio file format and stored in a database. The calls are selected and reviewed for any analysis or review.

2.2. Challenges

Some challenges with the current process, as shown in Figure 1 are:

- When calls are recorded in audio file formats, the ability to perform quantitative analysis diminishes. These calls require translation or language localization before being used for any analysis.
- Due to the impracticality of manually reviewing and analysing every recorded call, a selection process based on criteria such as call recency or nature is employed. Industry data indicates that less than 1% of calls undergo analysis and review. Consequently, this small sample size introduces bias in decision-making.

3. Materials and Methods

To enhance the efficiency of customer support executives and uncover valuable business trends and insights from recorded calls, we propose leveraging Google Cloud Storage (GCS) and Platform services.

The end application not only makes it easy for users with different technical expertise to interact with ML models and consume data insights but also holds the potential for uncovering valuable business insights that could significantly impact call center operations, as shown in Figure 3.

3.1. Proposed Solution

3.1.1. Storing Recorded Calls

Store recorded calls as raw data in GCS and build an ETL data pipeline to consume it in Big Query. This can be done using Cloud Data prep or object tables to access GCS data in Big Query.

3.1.2. Transcribing Calls

Use Speech-to-Text API to automatically transcribe calls. API supports 100+ languages.

3.1.3. Sentiment Analysis

Use the Cloud Natural Language Processing API ‘s analyze_sentiment
feature to perform sentiment analysis on transcribed calls.

3.1.4. Visualization

Integrate with Looker Studio for easy visualization and monitoring of key performance metrics.

3.1.5. BI Chatbot

Use the Large Language Model (LLM) to build a chatbot layer on top of Big Query metadata to answer key metrics-related questions.

- Accessing Audio Files in BigQuery: Object tables in BigQuery help process and query unstructured data records such as audio files without complex data pipelines. Audio files stored in GCS can be accessed in BigQuery using object tables. The code snippet given in Figure 4 can create object tables on unstructured records in GCS.
- Leveraging Speech-to-Text API: Cloud Speech-to-Text API on Vertex AI enables precise audio-to-text transcribing. Pre-trained models are useful for call centers as they provide accurate transcriptions with minimal AI/ML effort. The features in the model can be customized based on specific use cases. Fig.5 code snippet can be used to invoke a remote Speech-to-text model.
- Integrate NLP API for Sentiment Analysis: After transcribing calls, use NLP models in Big Query to perform text analysis and understand the sentiment of the calls. ‘Analyse_sentiment’ feature of the model performs sentiment analysis, as mentioned in Fig.6. The transcribed calls can be bucketed into 3 easy-to-understand categories - positive, negative, or neutral basis sentiment scores.
- Visualize ML Model Results in Looker Studio: Big Query data can be integrated with Looker Studio to create reporting dashboards.
- Example dashboards include:
  - Agent Performance: Metrics include the number of conversations, sentiment distribution, and overall sentiment score. This can be used to assess agent performance.
  - Conversation Overview: Metrics include total conversations, average call duration, and sentiment categorization. This can be used to assess overall call center performance and draw in-depth insights to improve efficiency.
  - Cloud and Topic Classification: Visualize keywords and topics discussed in calls to understand key discussion points.
- BI Chatbot: A chatbot layer, as shown in Figure 7, can be built on top of BigQuery metadata tables to answer simple analytics questions for executives. This is done using LLM models that understand the user prompt and fetch the desired result from underlying BigQuery data.
- Additional Use Cases:
  - Quality Scoring: Rank all audio calls and agents for quality checks using NLP models.
  - Forecasting Conversations: Use ML models to predict the number of future conversations, helping to improve preparedness and reduce wait times.
  - Personalized Responses: Generate personalized text responses for customers based on audio conversations.
  - NPS and CSAT Integration: Visualize Net Promoter Score and Customer Satisfaction Score alongside sentiment scores and agent performance metrics.

4. Results and Discussion

The proposed AI/ML architecture for analysing audio calls can improve visibility into trends, uncover more profound insights from data, and enhance overall call center efficiency. The solution, designed to be easily implemented within the Google ecosystem, requires minimal data movement and API integration. By leveraging these technologies, call centers can significantly improve their operations and customer service experiences, with the reassurance that the implementation process is straightforward and manageable.

4.1. Figure and Table

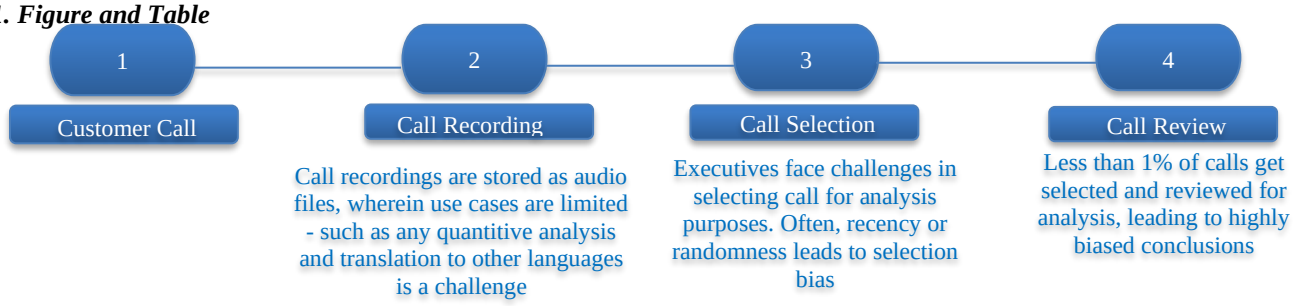

Fig. 1 Current call center process and challenges

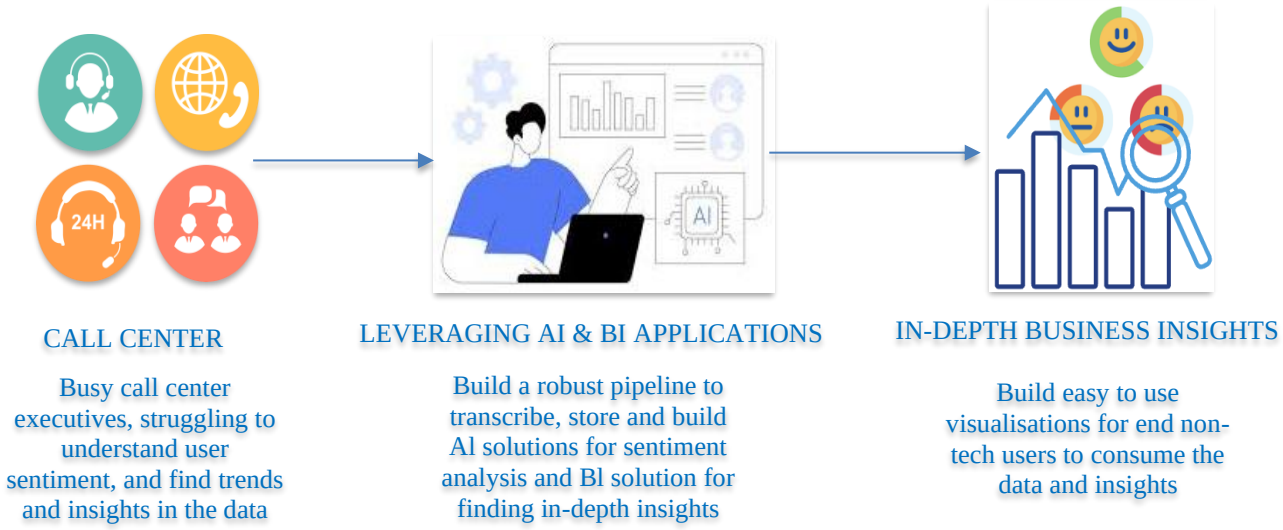

Fig. 2 Proposed solution

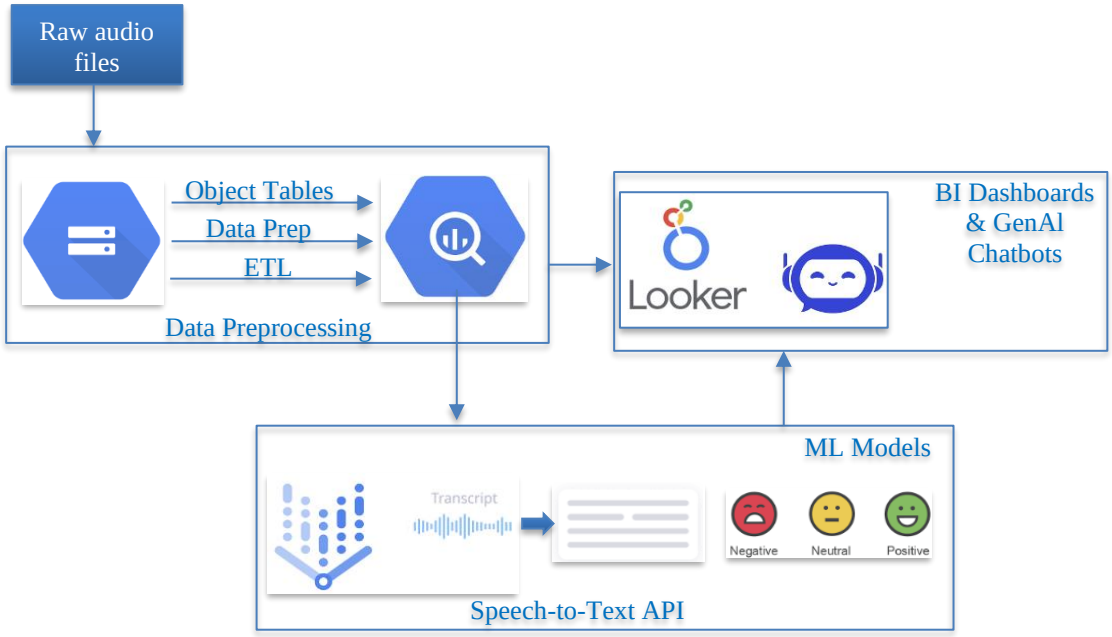

Fig. 3 Detailed solution architecture

```
--Create an object table
CREATE OR REPLACE EXTERNAL TABLE `name_of_table`
WITH CONNECTION `connection_name` --cloud resource connection
OPTIONS(uris=["gs://mybucket/audio/*"],  -- cloud locationwhere audio files are stored
        object_metadata="SIMPLE");
```

Fig. 4 Accessing audio files in BigQuery

```
CREATE OR REPLACE MODEL
`project_id.dataset.nlp_model_name` -- GCP project ID and dataset id wherein the model will reside
REMOTE WITH CONNECTION project_id.region.connection_id -- region and connection id of cloud resource connection
OPTIONS (REMOTE_SERVICE_TYPE = 'CLOUD_AI_SPEECH_TO_TEXT_V2')
```

```
SELECT * FROM
ML.TRANSCRIBE(
  MODEL `project_id.dataset.model_name`, --GCP project ID and remote model name
  TABLE `project_id.dataset.object_table`, -- object table with audio data (created in previous step)
  [RECOGNITION_CONFIG => ( JSON 'recognition_config')] --required when no recognizer has been specified for remote model.
)
```

Fig. 5 Remote Speech-to-text API Model Call in BigQuery

```
CREATE OR REPLACE MODEL
`project_id.dataset.nlp_model_name` -- GCP project ID and dataset id wherein the model will reside
REMOTE WITH CONNECTION project_id.region.connection_id -- region and connection id of cloud resource connection
OPTIONS (REMOTE_SERVICE_TYPE = 'CLOUD_AI_NATURAL_LANGUAGE_V1')
```

```
SELECT * FROM ML.UNDERSTAND_TEXT(
  MODEL `project_id.dataset.nlp_model_name`, --GCP project ID and remote NLP model name
  TABLE `table_name`, --Bigquery table where transcribed data is stored
  STRUCT('analyze_sentiment' AS nlu_option) --Feature name in NLP model
)
```

Fig. 6 Remote Sentiment Analysis Model Call in BigQuery

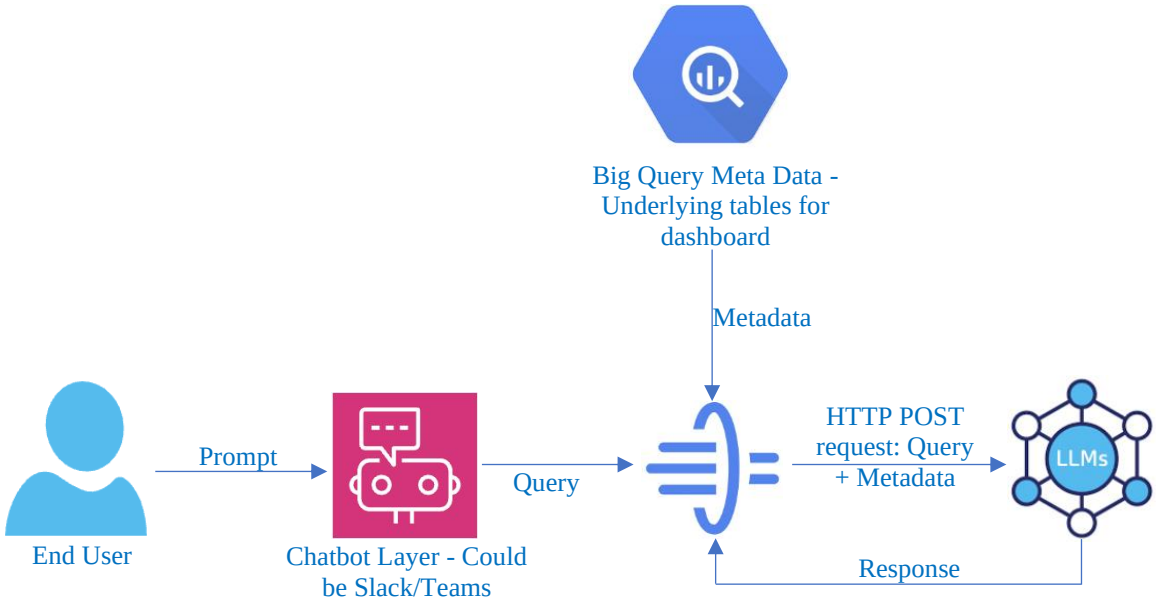

Fig.7 Overview of ChatBot for answering analytics questions

References

[1] Amazon, CCI Post Call Analytics Solution. [Online]. Available: https://aws.amazon.com/machine-learning/ml-use-cases/contact-center-intelligence/post-call-analytics-pca/

[2] Google Cloud, The ML. TRANSCRIBE Function. [Online]. Available: https://cloud.google.com/bigquery/docs/reference/standard-sql/bigqueryml-syntax-transcribe

[3] Google Cloud, The ML. UNDERSTAND_TEXT Function. [Online]. Available: https://cloud.google.com/bigquery/docs/reference/standard-sql/bigqueryml-syntax-understand-text

[4] Goutam Chakraborty, Murali Pagolu, and Satish Garla, *Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS*, SAS Institute, 2014. [Google Scholar]

[5] Awni Hannun et al., "DeepSpeech: Scaling up end-to-end Speech Recognition," *arXiv*, 2014. [CrossRef] [Google Scholar] [Publisher Link]