



COnVoy: A Contact Center Operated Pipeline for Voice of Customer Discovery

Rishabh Tripathi[†], Digvijay Anil Ingle[†], Ayush Kumar, Cijo George, Jithendra Vepa

Observe.AI, India

{rishabh.tripathi, digvijay.ingle, ayush, cijo.george, jithendra}@observe.ai

Abstract

With ever increasing volume of contact centers globally, there is a massive flow of customer service interactions into contact center telephony systems. These interactions are a rich source of information that can be utilized to uncover valuable insights regarding customer preferences and pain-points. In this paper, we propose an unsupervised pipeline comprising of an ASR engine, a reason-utterance detector, a re-writing module, a topic modelling block and a topic-description generator to enable *Voice of Customer (VoC) Discovery* by systematically tracking the reasons for the conversations. The insights surfaced by our pipeline can help contact centers get a bird's eye view of what is happening in their business. This would not only help business leaders come up with strategic measures to provide high quality and quicker resolutions to customers but also drive better customer satisfaction.

Index Terms: reason of call, contact center, speech analytics

1. Introduction

Contact centers play a vital role in addressing customer issues or queries and delivering positive customer experience. Large scale contact centers, typically handle millions of interactions daily. These interactions consist of a plethora of information capable of generating valuable insights into customer behaviour, preferences and pain-points. These insights can be further utilized to take strategic measures to improve the quality of customer service, operational efficiency and achieve better customer satisfaction levels. One of the most important aspects of effective contact center operations is the ability to identify and track the reasons for customer service interactions (refer Table 1). These conversations are typically characterized by:

Diversity in reason of the call: Customers may reach out to contact centers for various reasons such as, product or service enquires, technical support, feedback or complaints, etc. For example, in a typical e-commerce platform, the reasons for which a customer might reach out could range from as simple as "placing an order" to as complex as "reporting delivery of a duplicate phone". Hence, contact centers often watch out for the top reasons that their customers are reaching out for. The idea here is to surface the trends in call volume and identify opportunities to minimize these calls and assist agents in providing seamless resolution to such issues with minimum effort.

High dynamicity: Customer service environment is highly dynamic in nature. Most common issues that customer's reach out for in this week may not be the same in consequent week. For instance, if an e-commerce platform launches a large scale promotional campaign offering discounts on specific credit

Table 1: Representative example. Highlighted **Red text** represents utterance indicating reason of the call (Reason-Utterance)

Agent: good morning this is **** how are you doing
Customer: i am doing good i needed a help so i called
Agent: sure go ahead i am here to assist you
Customer: so i have a frozen account with you but i am not sure i have been charged this month for no reason and umm so i am looking for a refund here with you guys
Rewritten reason of the call: I want refund for an inaccurate charge raised on my frozen
Short Description: Refund for inaccurate charge

cards, it is quite natural to expect higher proportion of calls in association with the campaign as opposed to previous week. As a result, it is equally important for contact centers to surface newly emerging issues that would otherwise be far from their line of sight. Tracking them would help contact centers update their knowledge base to equip the agents in providing quick and concise resolution to their customers for such issues.

In this work, we demonstrate an end-to-end pipeline to enable *VoC Discovery* for contact centers by:

1. Detecting the reason for an interaction and consequently generating an executive summary to provide bird's eye view of the business over a given period.
2. Surfacing newly emerging and out-of-distribution reasons for which customers are calling.
3. Ability to map the reason of the call to evidence in the conversation so as to enable further investigation into conversation.

2. Pipeline And Components

2.1. Automatic Speech Recognition Engine (ASR-Engine)

We utilize a third-party ASR Engine to transcribe audio conversations into text. The calls comprise of English dyadic conversations belonging to a variety of industries like, healthcare, banking *etc.* We observe a Word Error Rate of 15-20% on calls belonging to these industries.

2.2. Reason-Utterance Detector

Reason-utterance detector block consists of a text classifier aimed at predicting whether a given turn in the text transcript represents the reason for the call or not (eg: highlighted turn in Table 1). We fine-tune DistilRoBERTa-base¹ on proprietary conversational dataset. The accuracy of the classifier on the test set is 88.23%.

[†]Equal contribution

¹<https://huggingface.co/distilroberta-base>

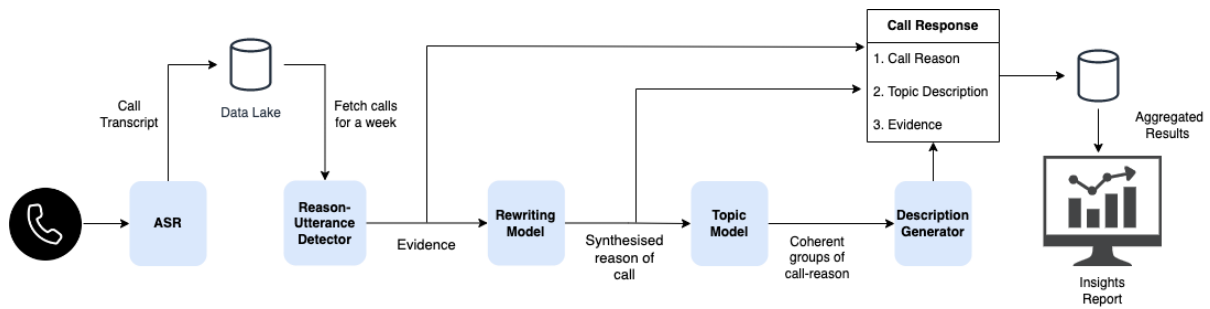


Figure 1: Architecture Diagram

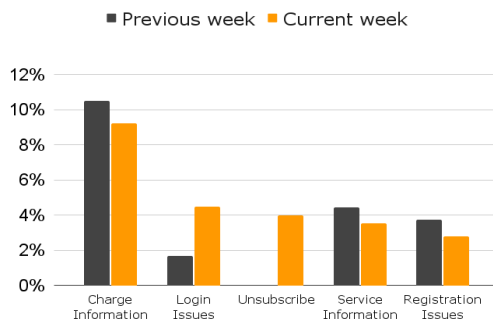


Figure 2: Weekly analysis of top-5 reasons of the call for a business account for the first two weeks in the month of March

2.3. Rewriting module

The reason-utterance detected in the previous step is an excerpt from the actual agent-customer spoken conversation which comprises of several nuances such as ASR errors, disfluencies, co-references, etc. In order to de-contextualize it, we couple the turn indicative of call-reason with surrounding context and use our in-house T5-based [1] large language model (LLM) to generate a *synthesized* version of the reason of the call.

2.4. Topic Modelling

We perform topic modeling over the synthesized call-reasons using BERTopic [2] to cluster them into coherent groups. We observe a coherence score of 0.53 on a proprietary test set.

2.5. Describing the coherent groups

We utilize our in-house language T5-based model that has also been trained to generate a short description (eg: *Charge Information*, *Registration Issues* in Figure 2) when given a list of representative examples belonging to a call driver group. This short description helps business stakeholders in getting a quick understanding of the data.

3. Voice of the Customer Discovery

Contact centers often have weekly business reviews that allow business leaders to evaluate the performance of contact center in terms of key-metrics, such as call volume, call-handling time, customer satisfaction score, and other relevant performance indicators. They help identify areas of improvement and address any issues that may arise during the week. Our pipeline for discovering *VoC* aids these reviews by providing valuable insights

and makes sure that the actions are backed by data.

Every incoming call is transcribed using the ASR engine and ingested into a central data lake. Our pipeline functions in a batch-processing manner and is triggered on a weekly basis to process the call transcripts belonging to only one week. Further, it ingests the synthesized reason of the call, topic description and corresponding evidence in a database that is streamed to an interactive dashboard in the form of an *Insights Report* based on user requests. Specifically, the report provides a summary of the call-volume distributed across the themes identified by our pipeline to compare and contrast with the distribution in previous week. An illustrative example of such a summary is presented in Figure 2. We observe that the theme *Login Issues* has seen almost twice increase in the call volume compared to previous week. This might be insightful to contact centers as they might want to reach out to their technology teams to investigate for the cause of this and decide on an appropriate course of actions. Furthermore, we also observe a newly emerging category *Unsubscribe* in the current week that was not present in the last week. Implying that the customers are requesting to get unsubscribed from their service which might be critical to their business. Hence, they would want to further deep dive into its root cause. Additionally, our proposed pipeline also provides a capability to further drill down into the evidences in the call-transcripts to better understand the context around these issues.

4. Conclusions

In this paper, we showcase an end-to-end pipeline to enable *VoC Discovery* for contact center conversations that can surface key-insights around specific reasons for which their customers are reaching out. It can also surface newly emerging themes of customer issues, thus helping contact centers in reducing the time-to-action to employ strategic measures to fix them. Furthermore, it can map the identified call-reasons to specific evidences in the conversation thus empowering contact centers to perform deeper analysis wherever necessary. In short, the insights surfaced by our pipeline can be used by contact centers to take strategic measures to improve the quality of service, operational efficiency and thus drive better customer satisfaction.

5. References

- [1] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *J. Mach. Learn. Res.*, vol. 21, pp. 140:1–140:67, 2020.
- [2] M. Grootendorst, "Bertopic: Neural topic modeling with a class-based TF-IDF procedure," *CoRR*, vol. abs/2203.05794, 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2203.05794>