# Financial Customer Inquiries Project

TELUS AI Insights | CMC Customer and Channel Insight

Presented by Farhad Momeni Farahani
February 19th, 2025

# Problem Statement

- High volume of customer inquiries requiring efficient handling.
- Unstructured text data in customer interactions.
- No automated categorization, leading to inefficiencies.
- Need for data-driven insights to improve service.

# Agenda

01

## Key Insights

- Dataset with the size of 157,229
- Include customer interactions in the banking industry

02

## Approach and Methodology

- Data Preprocessing
- Embedding Generation
- Clustering
- Keyword Extraction

03

## Model Performance

- Classification Models
- Hyperparameter Tuning
- Best Performing Model

04

## Challenges and recommendations

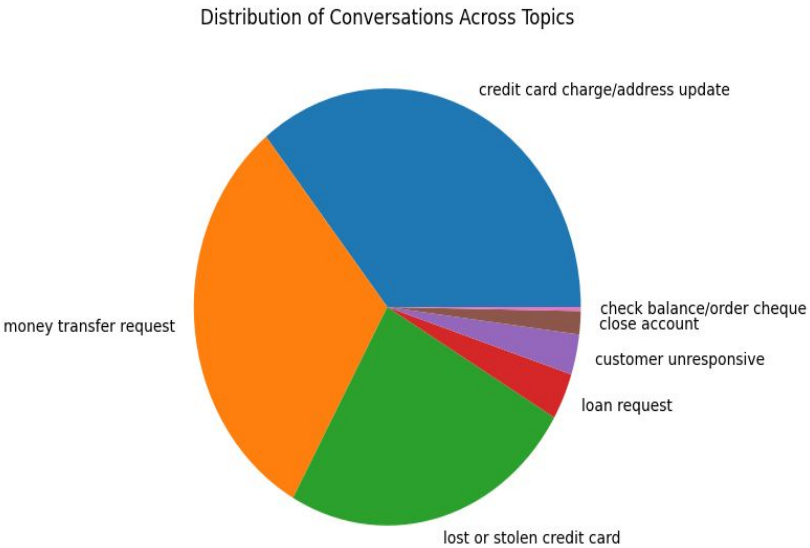
- High Similarity Between Inquiries
- Training Time for Large Embeddings

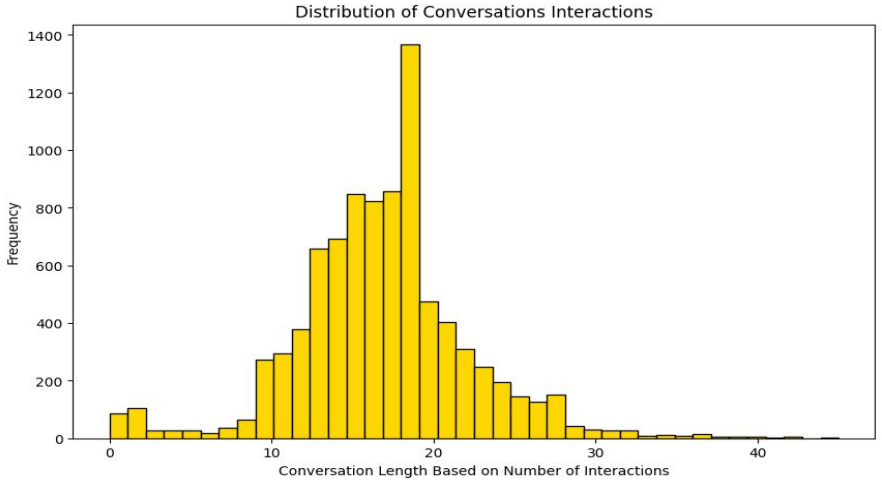
# Key Insights

- Customer support receives a high volume of recurring inquiries, showing opportunities for automation.
- Customer inquiries can be categorized into 7 key topics, which can help optimize response workflows.
- Many inquiries share common themes, meaning related topics could be grouped or automated more efficiently.
- Certain topics require more agent attention, while others could be handled with self-service solutions.
- Dimensionality reduction helps refine clustering and improves topic separation.

# Approach and Methodology

- Some inquiries require multiple interactions, while others are resolved quickly.
- Applied text preprocessing for having better word embedding.




- SBERT embeddings (384D) were used to convert text into numerical representations.
- Applied DBSCAN with UMAP (10D reduction) to identify key inquiry topics.
- Labeling clusters by reviewing top conversations with keyword extraction with TF-IDF.

# Model Performance

| Model/Score | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|
| Logistic Regression | 0.9916 | 0.98910 | 0.9891 | 0.9891 |
| SVC | 0.9935 | 0.99140 | 0.9913 | 0.9914 |
| Random Forest | 0.9814 | 0.9750 | 0.9782 | 0.9757 |
| LightGBM | 0.9874 | 0.9851 | 0.9850 | 0.9847 |
| XGBoost | 0.9818 | 0.9817 | 0.9818 | 0.9813 |
| CatBoost | 0.9850 | 0.9856 | 0.9855 | 0.9853 |
| MLP | 0.9895 | 0.9894 | 0.9895 | 0.9894 |
| DNN | 0.9859 | 0.9861 | 0.9859 | 0.9858 |

# Challenges and Limitations

- Computational limitations  for training and fine-tuning models.
- High similarity between some inquiries.
- Imbalanced data distribution.

# Business Recommendations

- Invest in scalable computing resources.
- Use chatbots and self-service portals for high-volume inquiries to reduce agent workload.
- Implement Real-Time Sentiment Monitoring.