Scotiabank

Chatbot for Financial Data

Scotiabank Quarterly Reports: Instant Insights with Custom Chatbot

Team Alpha

Our Team

Uma Maheshwari Balaji

Sonal Parmar

Raemil Corniel

Jharana Adhikari

Aparna Suresh

Agenda

- Introduction
- Data Extraction and Preprocessing
- Models Overview
- Evaluation Metrics
- Results
- Discussion
- Conclusion and Future Work
- References

Introduction - Problem Statement and Metrics

Problem Statement:

- Scotiabank's existing systems fall short in providing users with an efficient way to quickly interpret and interact with the complex financial data found in quarterly reports, resulting in delayed insights and diminished user satisfaction.

Significance:

This enhancement will provide users with immediate access to detailed financial information, improving their ability to make informed decisions and increasing engagement with Scotiabank's financial data.

Home Financial Analysis

Introduction - Project Overview

Objective:

- **Enhance Scotiabank's Chatbot:** Transform Scotiabank's quarterly reports into instant insights with a custom chatbot, designed to handle and respond to financial queries and metrics effectively.

Metrics:

- **Success Indicators:**
  - **Accuracy of Responses:** Measured by ROUGE and BLEU to assess how well the chatbot's answers match reference data.
  - **Quality of Summaries:** Evaluated using METEOR and Perplexity to ensure summaries are coherent and comprehensive.
  - **User Interaction Efficiency:** Assessed by how effectively the chatbot handles queries and provides timely, relevant information.

Data Preprocessing

Text Extraction:

- **Tools Used:** tabula-py, pypdf, pdfminer
- **Process:** Extracted text from Scotiabank's Q2 2024 PDF report.
- Analyzed 11 pages; extracted 7,373 words.
- 10 Tables

Preprocessing:

- **Cleaning:** Tokenization, normalization, removal of special characters, and handling of text inconsistencies.

Data Preprocessing: Challenges and Solutions

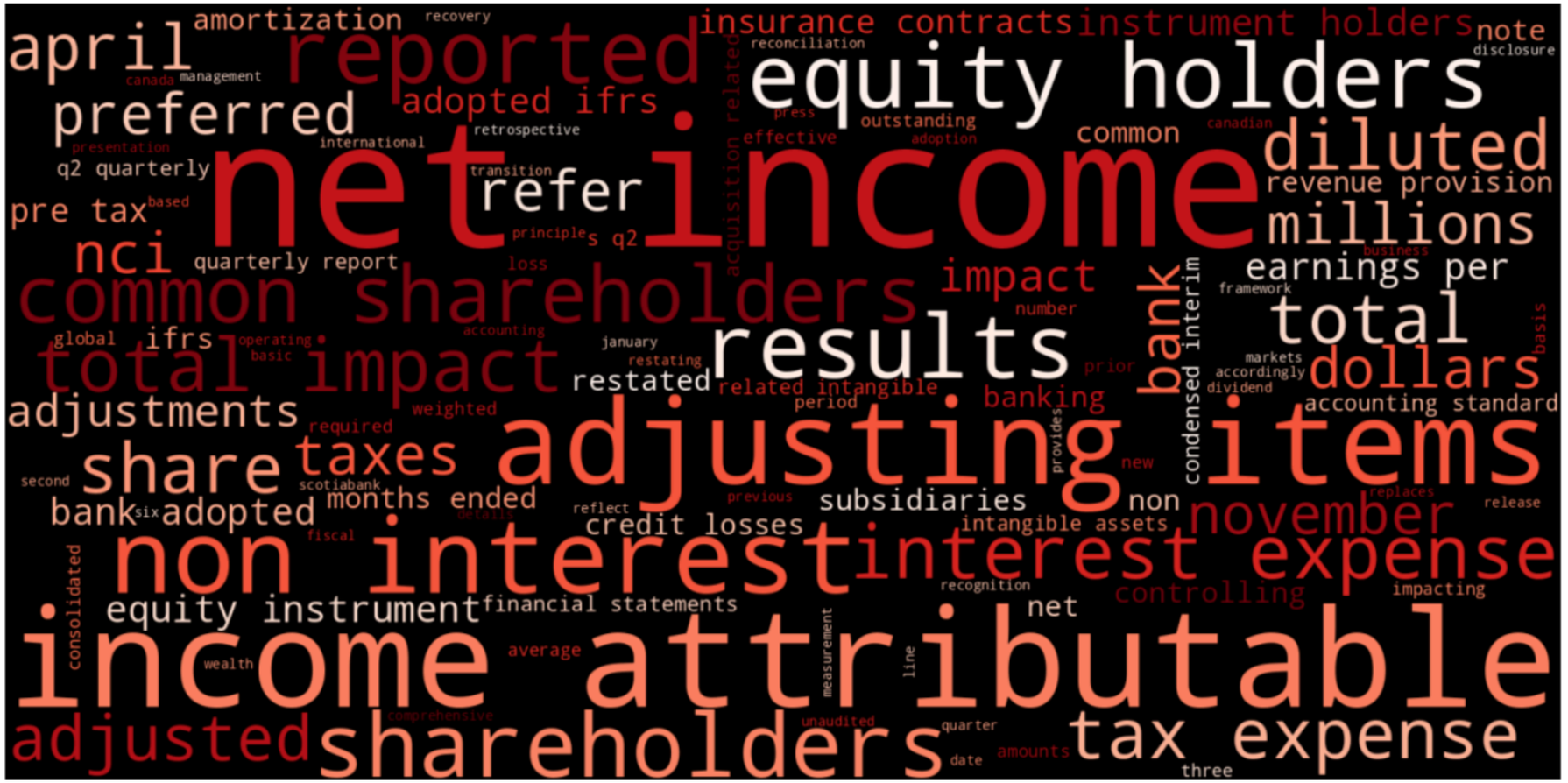
Merged Tables:

- **Issue:** Two tables that are in single page were merged.
- **Solution:** Used df.loc to index and slice the DataFrame, effectively dividing the merged table into separate, distinct tables.

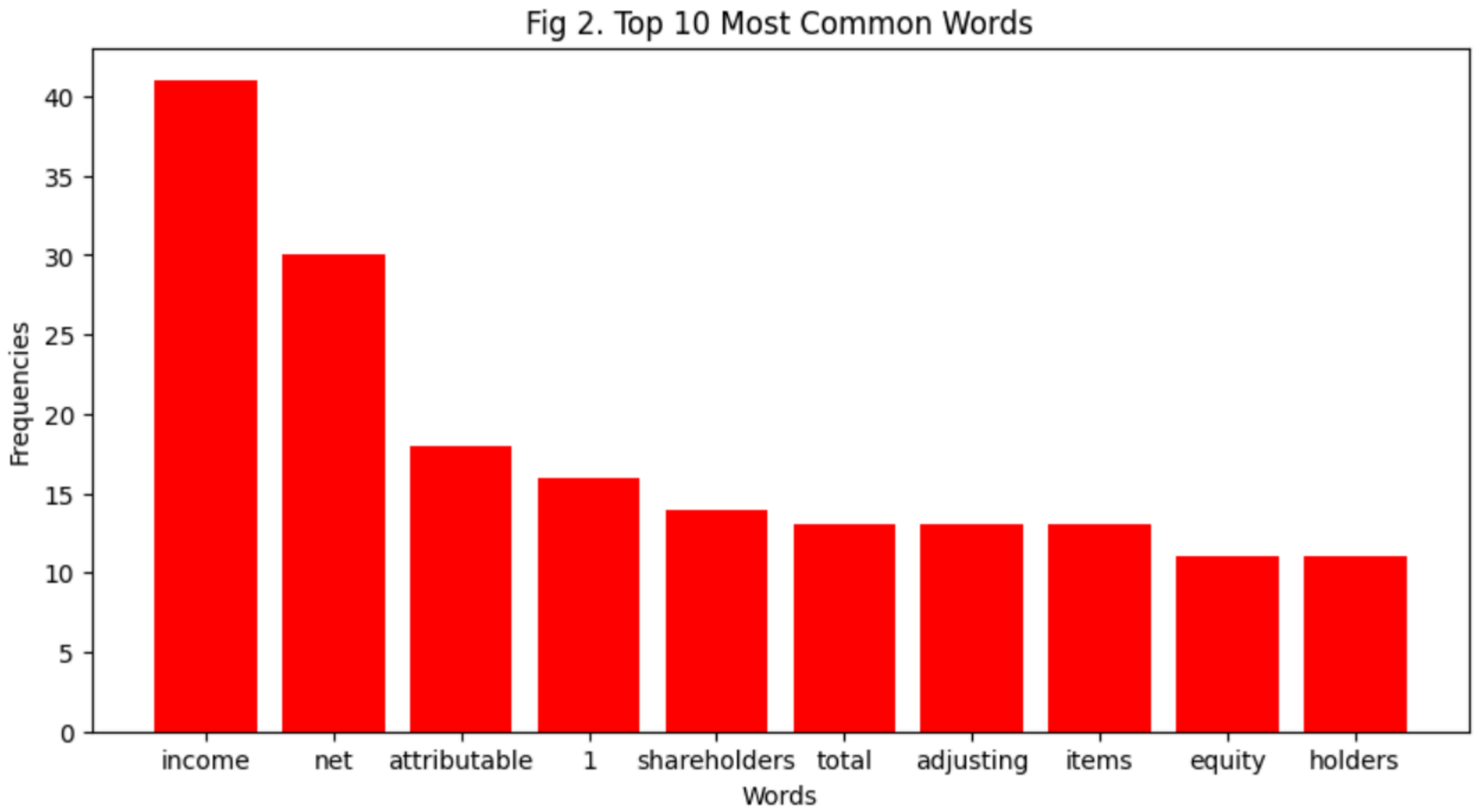
Merged Columns:

- **Issue:** Numeric and text data were combined in a few columns.
- **Solution:**
  - Applied regex to separate columns into numeric and text data.
  - Used space-based splitting for further column separation.

Exploratory Data Analysis (EDA):



Exploratory Data Analysis (EDA):



Exploratory Data Analysis (EDA):

Fig 3. Top 10 Bigrams

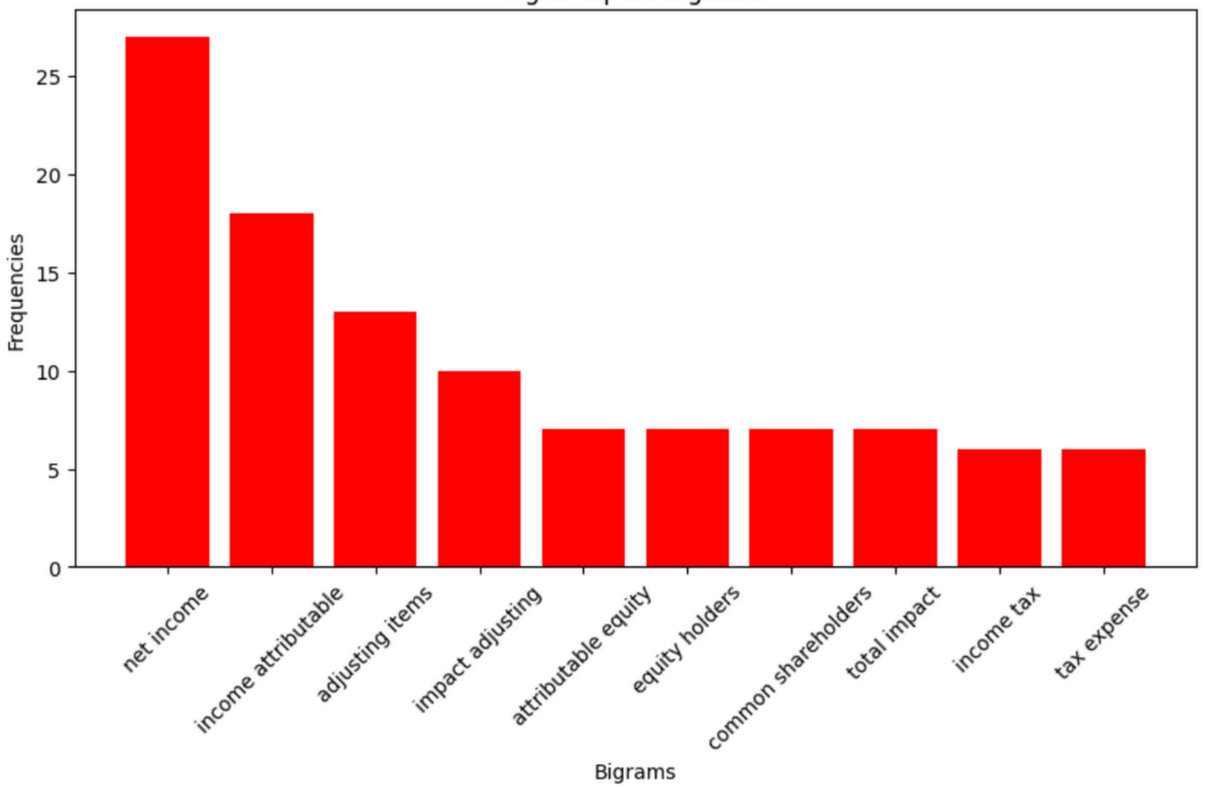
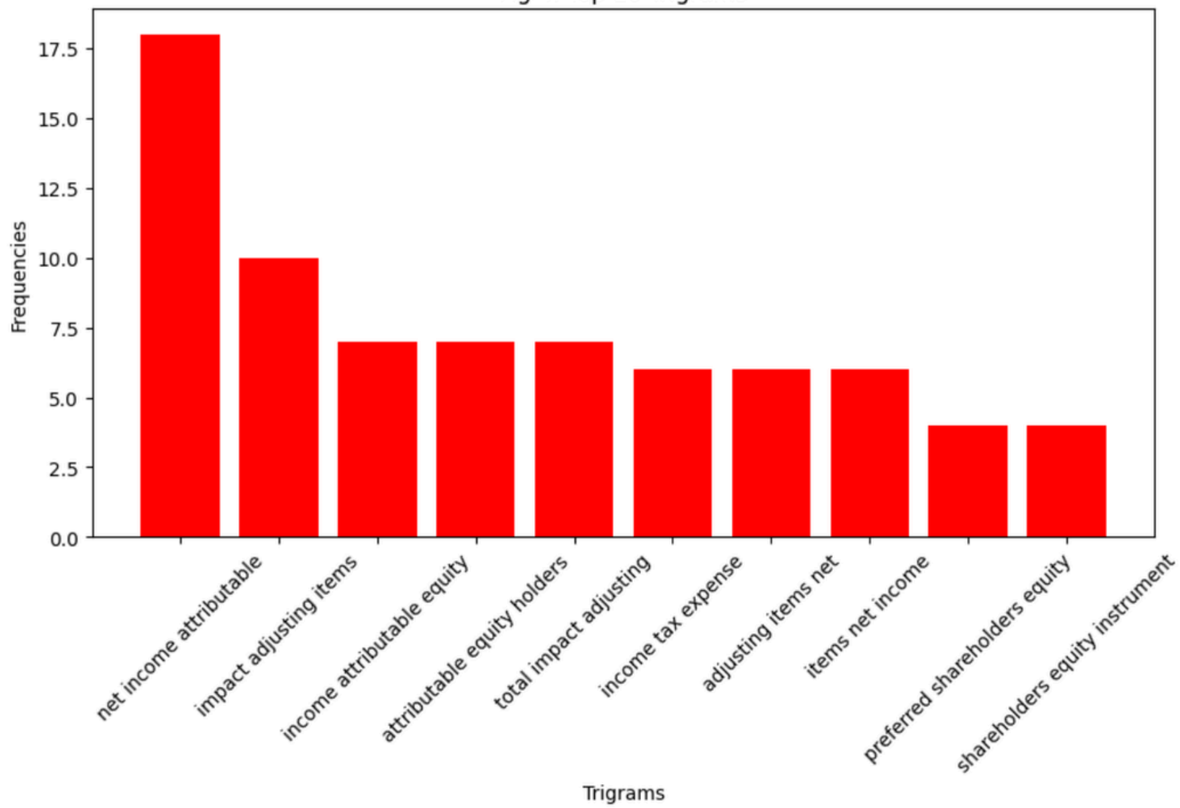


Fig 4. Top 10 Trigrams



Models Overview

Pipeline: "question-answering

Purpose:

To provide an easy-to-use interface that enables answering questions directly from a given text or context, streamlining the process of querying information from documents or datasets.

BertForQuestion Answering

Purpose:

To accurately extract answers from text by understanding the relationship between the question and the context, making it ideal for applications where precise information retrieval and language comprehension are critical

GPT2LMHeadModel

Purpose:

To generate coherent and contextually appropriate text based on a provided prompt, making it useful for tasks like completing sentences, generating creative content, or simulating human-like conversation.

Evaluation Metrics

ROUGE (Recall-Oriented Understudy for Gisting Evaluation):

Function: Measures overlap between generated summaries and reference text.

METEOR (Metric for Evaluation of Translation with Explicit ORdering):

Function: Considers synonyms, paraphrasing, and stemming to evaluate summary quality.

Perplexity:

Function: Measures how well a probability model predicts a sample. Lower perplexity indicates better model performance.

BLEU (Bilingual Evaluation Understudy Score)

Function: Measures matches between generated text and reference text to assess summary quality.

Pipeline: "question-answering

Account Management:

Assisting users with inquiries about account balances, recent transactions, and account status updates.

Answering questions:

Answering questions related to financial metrics that are based on Q2 report.

Investment Advice:

Providing information on investment options, portfolio performance, and market trends based on user queries

Successful activities

Enhanced user experience with strategies based on user financial goals and queries.

GPT2LMHeadModel

Information Extraction:

Excelled in producing concise and coherent summaries by extracting the financial information.

Readability:

Made complex financial data more accessible and easier to understand.

Comparison

Generated information with clear and straightforward language

Successful activities

Enhanced user experience with brief and digestible insights.

BertForQuestion Answering SentenceTransformer('all-MiniLM-L6-v2')

Contextual Understanding:

Demonstrated exceptional ability to interpret and understand financial text within the Flask app.

Insight Depth:

Provided comprehensive insights into the intricacies of the financial report, seamlessly integrated into the Flask interface.

Accuracy:

Ensured high accuracy in contextualizing complex financial information, as validated through real-time interactions.

Relevance:

Effectively identified and highlighted key details relevant to the financial content displayed in the Flask app.

Metrics

Metrics	Pipeline: "question-answering"	BertForQuestionAnswering	GPT2LMHeadModel
METEOR	0.9803644613233463	0.9995970990808904	0.9997985900982216
BLEU	0.9992750120504148	0.9971469038809805	0.985935162285558
ROUGE	rouge1: Score(precision=0.9713114754098361, recall=0.9957983193277311, fmeasure=0.9834024896265561) rouge2: Score(precision=0.9711934156378601, recall=0.9957805907172996, fmeasure=0.9833333333333334) rougeL: Score(precision=0.9713114754098361, recall=0.9957983193277311, fmeasure=0.9834024896265561)	rouge1: Score(precision=0.9957983193277311, recall=0.9957983193277311, fmeasure=0.9957983193277311) rouge2: Score(precision=0.9957805907172996, recall=0.9957805907172996, fmeasure=0.9957805907172996) rougeL: Score(precision=0.9957983193277311, recall=0.9957983193277311, fmeasure=0.99579831932773 11)	'rouge1': Score(precision=0.9754098360655737, recall=1.0, fmeasure=0.987551867219917 1) 'rouge2': Score(precision=0.9753086419753086, recall=1.0, fmeasure=0.9875) 'rougeL': Score(precision=0.9754098360655737, recall=1.0, fmeasure=0.987551867219917 1)
Perplexity			1.8028159879508027e+34

Scotiabank

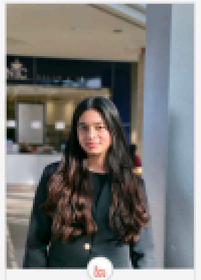
Home Financial Analysis About Us

AI for Financial Insights

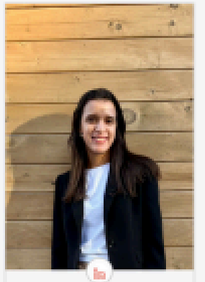
Financial insights is a vital component of overall financial analysis, yet many people struggle to access the support they need to manage their finances effectively. With the power of artificial intelligence and advanced data analysis, we can offer accessible and scalable financial support. This technology empowers individuals to make informed financial decisions, providing personalized guidance and insights to enhance their financial stability and growth.

Team Alpha

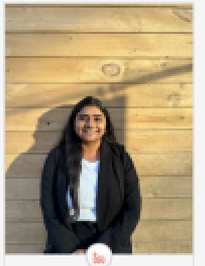
Meet our dedicated team of experts, passionately working behind the scenes to create an innovative financial insights chatbot designed to support and guide you on your journey to financial analysis and translations.



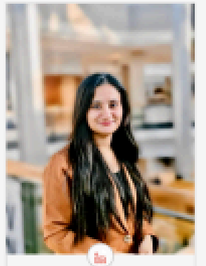
Uma Maheshwari



Raemil Cornel



Sonal Parmar



Jharana Adhikari

Deployed the Chatbot Using Flask API

Scotiabank

Home F Financial Anal

Objective

Develop an NLP-based chatbot to provide insights about Scotiabank quarterly financial results.

Methodology

A wide range of NLP techniques are utilized to handle user inputs and provide more accurate responses.

Design

A simple and user-friendly design that processes user input and works on a knowledge base to provide the best user feedback

Financial Insights Assistant

11:05
income tax expense

Financial Bot
11:05
total(2) reported net income (loss) $ 1,008 $

Financial Data

Relevant data from Table 1:

(Unaudited) ($ millions)	April 30 2024	January 31 2024	April 30 2023	April 30 2024.1	April 30 2023.1
income tax expense	$537	$533	$484	$1,070	$1,589

Send a message...

Choose File No file chosen

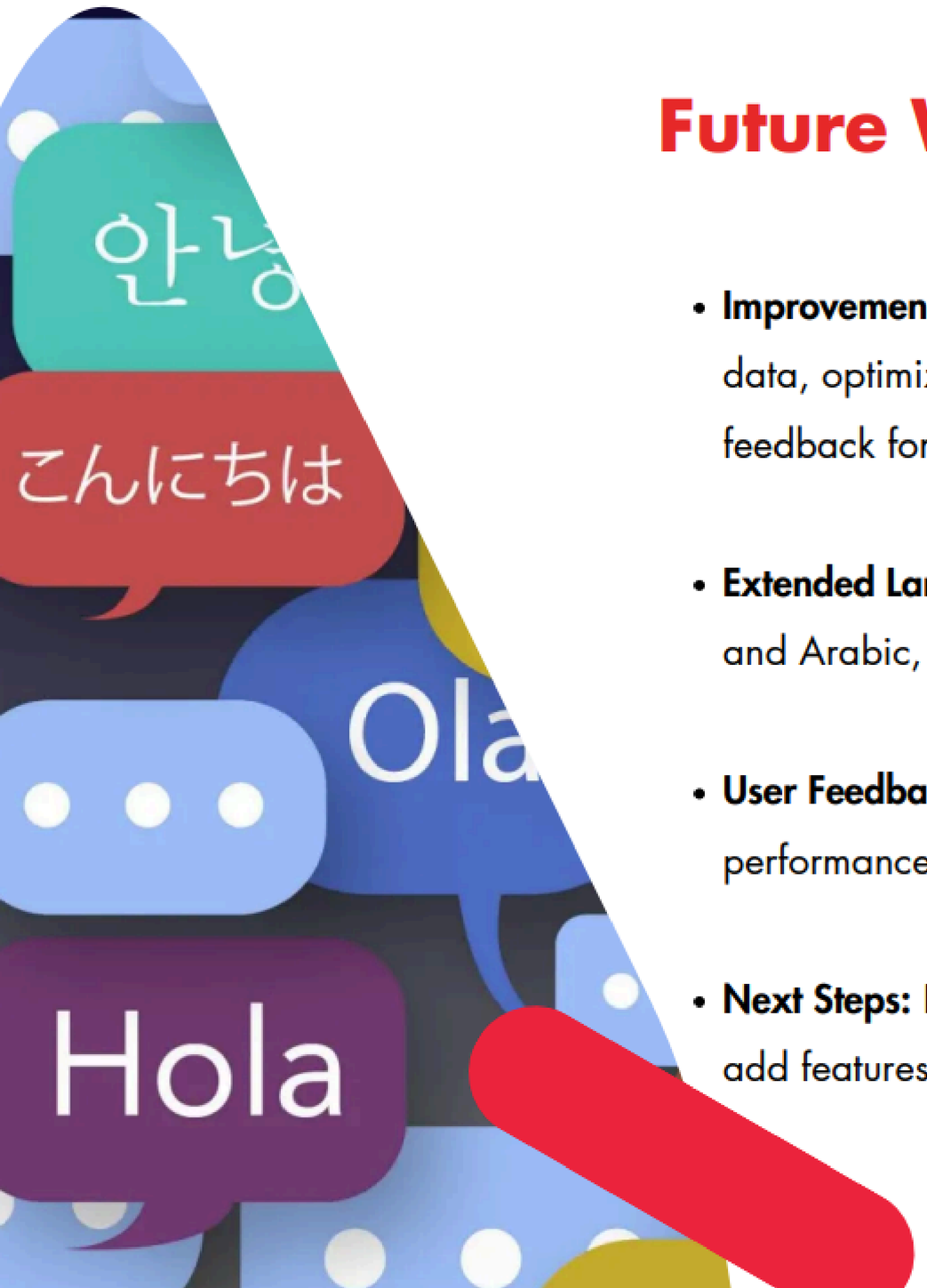
Conclusion:

Achievements:

- Extracted and summarized Scotiabank's Q2 2024 financial data and integrated it into a custom chatbot.
- Utilized precise techniques for accurate data handling, enhancing chatbot performance.

Contributions:

- Improved user experience by enabling the chatbot to deliver quick insights and key metrics.
- Increased satisfaction and engagement through effective summarization of complex financial information.



Future Work:

- **Improvements:** Enhance chatbot performance by fine-tuning models with domain-specific data, optimizing algorithms for better context understanding, and incorporating user feedback for refined responses.
- **Extended Language Support:** Broaden language capabilities to include German, French, and Arabic, expanding reach to a global audience.
- **User Feedback Mechanism:** Integrate user feedback to continuously improve model performance and response accuracy.
- **Next Steps:** Explore advanced models like GPT-4, incorporate multimodal capabilities, and add features such as real-time data integration and personalized recommendations.

Refrerencs

- https://www.scotiabank.com/content/dam/scotiabank/corporate/quarterlyreports/2024/q2/Q224_Quarterly_Press_Release-EN.pdf Li, Y., Wang, S., Ding, H., & Chen, H. (2024). Large Language Models in Finance: A Survey. arXiv:2311.10723v2.
- https://arxiv.org/abs/2311.10723 Ausekar, A., & Bhagwat, R. (2024). Banking on AI: Exploring the Transformative Role of ChatGPT in Financial Services. [Journal Name,
- Volume(Issue), Pages]. https://www.example.com Pillai, D., Doifode, A., Singh, A. S., & Bhosale, T. (2024). A Perspective on the Application of Fintech as a Gateway for Financial
- Literacy. [Journal Name, Volume(Issue), Pages]. https://www.example.com Thodge, A., M, H., & Cyril, P. D. (2024).
- Web-Based Chatbot for Basic Financial and Mortgage Services. [Journal Name, Volume(Issue), Pages]. https://www.example.com

Scotiabank

Thank You

Team Alpha