رواد مصر الرقمية

Customer Feedback Analysis

Ahmed Gamel   Alaa Awad

Ganna Zayed   Rowyda Abdelrahem

Hams Tamer   Salma Mohsen

# CUSTOMER FEEDBACK ANALYSIS AND IMPROVEMENT

Introduction

This project focuses on analyzing customer feedback to extract valuable insights, improve services, and enhance customer satisfaction.

The workflow is structured into several key stages, each contributing to a holistic solution for data analysis and improved decision-making.

It begins with setting up an SQL database to efficiently store and organize customer data, followed by designing a data warehouse that systematically aggregates and manages the data for advanced analysis.

In the next stage, sentiment analysis is applied to customer feedback, identifying positive and negative sentiments to help the business understand customer perspectives more deeply. Finally, the model is deployed using modern data science tools, operating in real-time and enabling immediate interaction with the data. This feedback loop fosters continuous improvement based on real-time customer responses. The project's ultimate goal is to provide an AI-powered platform that automatically and accurately analyzes customer feedback, leading to ongoing enhancements in service quality and operational efficiency.

SQL DATABASE SETUP AND DATA COLLECTION
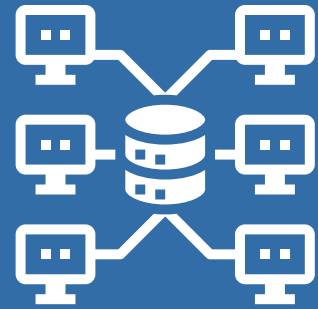
The first phase of the project involved setting up an SQL server using Microsoft Azure, which allowed the team to collaborate efficiently. The database was designed with tables for storing customer feedback forms and user data. Historical feedback data, containing 568,454 rows and 9 columns (Id, ProductId, UserId, Helpfulness Numerator, Helpfulness Denominator, Score, Time, Summary, Text), was imported into the database for analysis.

SQL queries were written to extract, summarize, and manipulate the feedback data, providing a comprehensive view of customer sentiments and behavior.

Microsoft SQL Server (Azure Cloud) and SQL Management Studio were the key tools used to manage the database, ensuring smooth data handling and team collaboration.

DATA WAREHOUSE AND PYTHON DATA PROCESSING
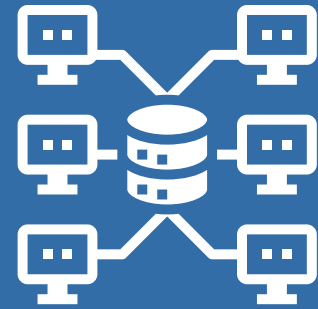
To facilitate in-depth analysis, a data warehouse was implemented. A Fact table named FeedbackFact and two dimension tables **(TimeDimension and UsersDimension)** were created. Fact table **(FeedbackFact)** that has quantitative data and metrics related to **feedback:feedback_id,product_id,user_id,helpfulness_numerator, helpfullness_denominator,score,time_id,summary,review_text.** First dimention table is TimeDimention that has **(time_id,feedback_date,year,month,day).**Second one is UsersDimention that has **(user_id, profile_name).** They store user related information which allows to analyze feedback. Feedback data was integrated into the data warehouse using Microsoft SQL Data Warehouse. Python, with libraries like Pandas and NLTK, was used to clean and preprocess the data for sentiment analysis. Data ingestion from the SQL server was accomplished via the pymssql library.
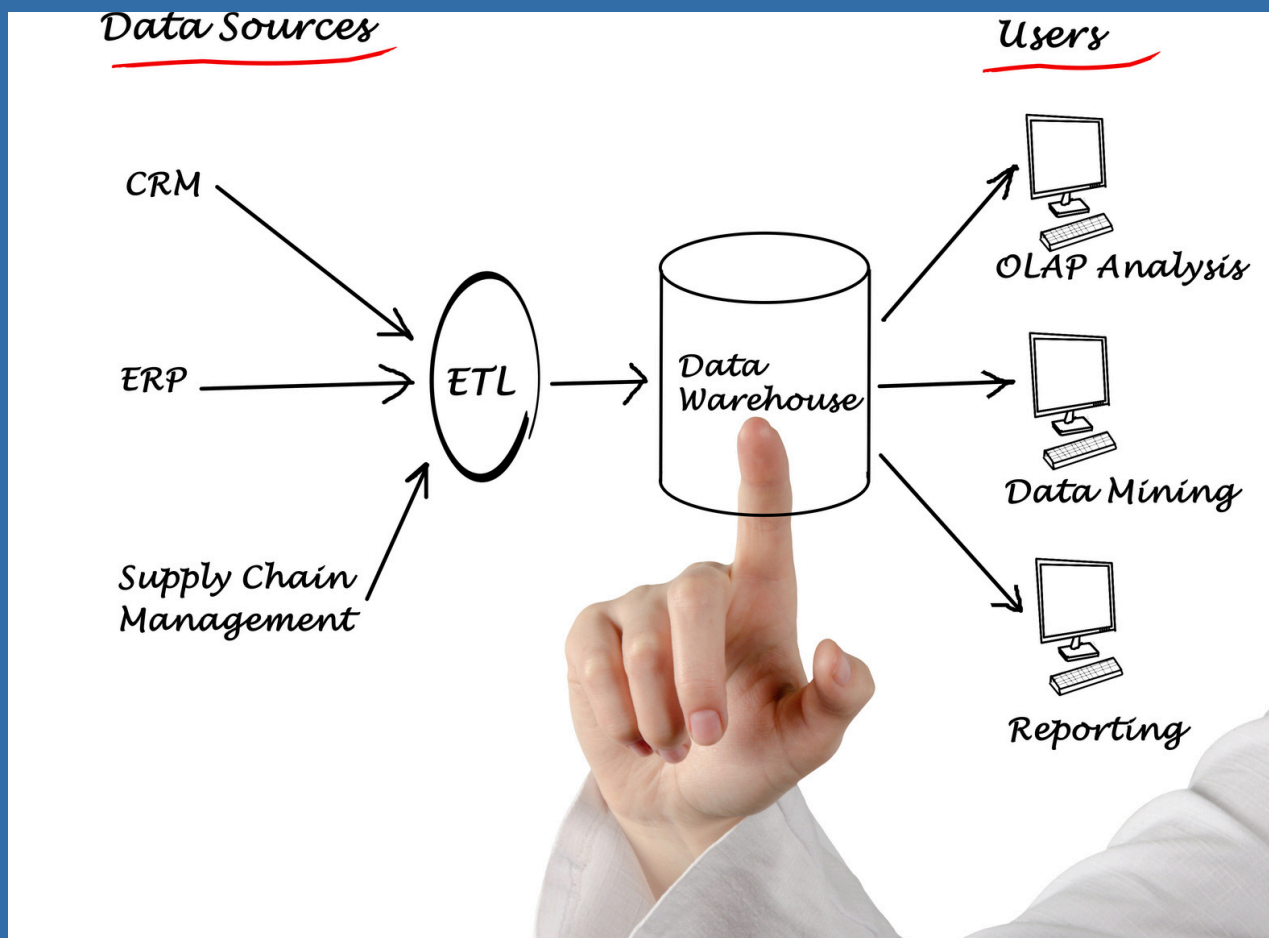
Model Training Metrics Over Epochs
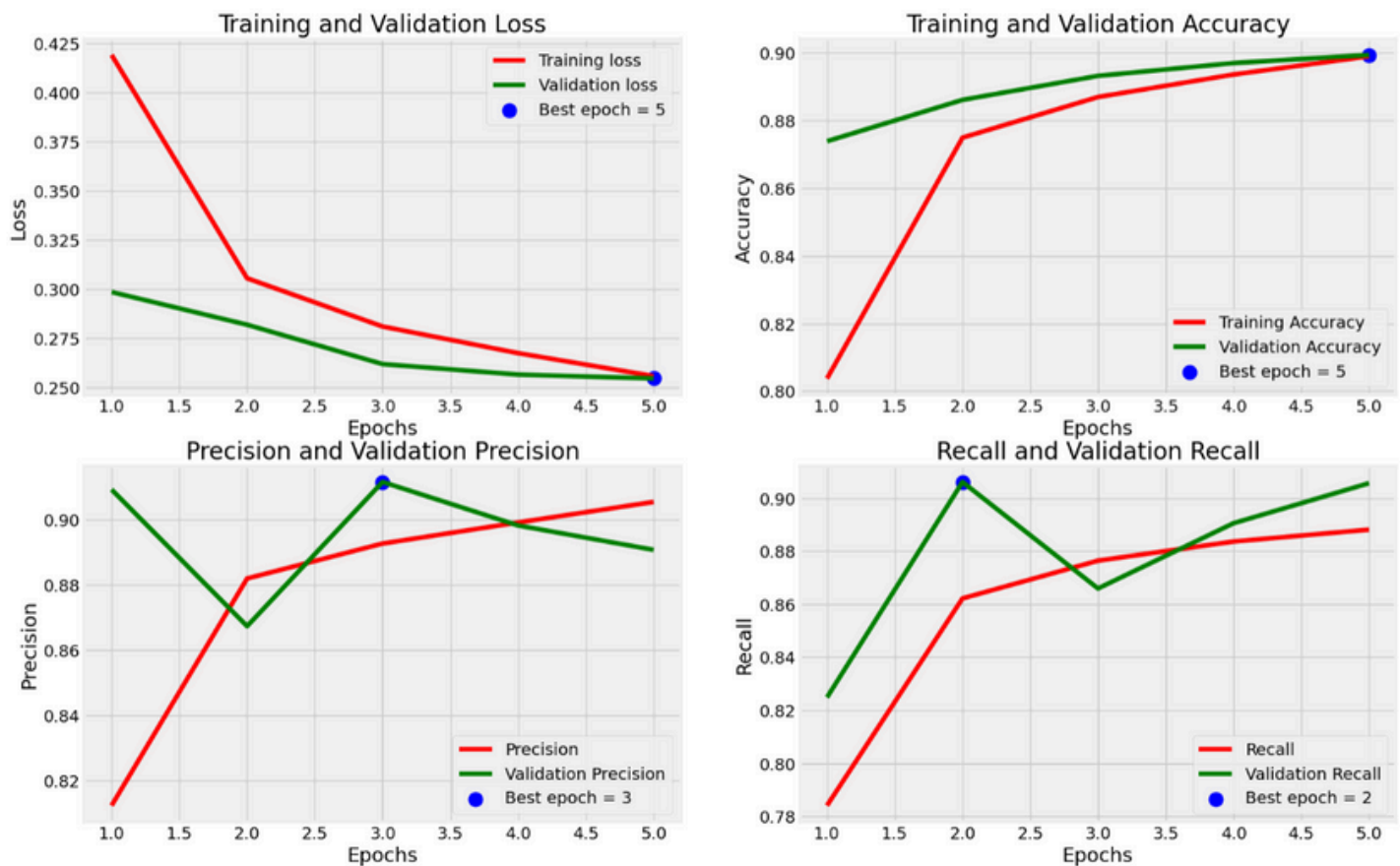
SENTIMENT ANALYSIS MODEL

In the sentiment analysis stage, a machine learning model was developed using Python to classify customer feedback as either positive or negative. The dataset, consisting of 162,037 rows with features like Helpfulness Numerator, Score, Summary, and Text, was used to train the model. A Long Short-Term Memory (LSTM) neural network was chosen for its effectiveness in processing sequential text data. The model architecture featured two Bidirectional LSTM layers to better capture the context of customer feedback, along with an embedding layer for word vector representations. Dropout layers were introduced to prevent overfitting, ensuring that the model generalizes well to new, unseen data.

Model Performance
The model was trained using a dataset split into three subsets: 110,000 samples for training, 27,037 for validation, and 25,000 for testing. After 5 epochs of training, the model reached a validation accuracy of 89.93%. Precision was recorded at 90.93%, and recall at 82.49%, indicating that the model performed well in identifying both positive and negative feedback.

MODEL DEPLOYMENT

The sentiment analysis model was deployed using Streamlit and integrated with Azure services to deliver a user-friendly and efficient experience. The web application allows users to input customer feedback directly, where the model analyzes the text and classifies it as either positive or negative, while displaying a confidence percentage for the prediction. The application was developed and deployed via Streamlit, offering a real-time interactive interface. By leveraging Azure's cloud infrastructure, the model operates in a secure, scalable, and highly available environment. Azure's monitoring tools are also utilized to track application performance, usage metrics, and any potential issues, ensuring optimal operation and user experience.The application can be accessed via the following link:
https://customer-project.azurewebsites.net/

CONCLUSION

This project demonstrated a successful approach to analyzing and improving customer feedback using state-of-the-art data science and machine learning methods. The SQL database provided a structured system for managing large datasets, while the data warehouse facilitated in-depth analysis. The LSTM model achieved high accuracy in classifying customer feedback, and its deployment on Azure ensured real-time analysis and scalability. By continuously analyzing customer feedback, the platform empowers businesses to enhance their services and make data-driven decisions that improve customer satisfaction over time.

Meet Our Best Team

Ahmed Gamel

Alaa Awad

Rowyda Abdelrahem

Ganna Zayed

Hams Tamer

Salma Mohsen

thank you