

System Pipeline Sequence

(Updated Version of Pipeline Sequence is shown in the presentation)

Data Store

Azure SQL Database

- Time series data (stocks, indexes, etc)

Blob Storage

- Question-Answer Pairs in Json
- Azure Data Factory or Azure Databricks for preprocessing

Blob Storage

- FAQ and answers (raw data)

Knowledge Base

- Question and answer pairs
- Supports QnA Maker API for queries

Azure SQL Database

- User Authentication Info
- User Data (Store with user permission)

Prepare Data

Data Preparation

1. Normalization
2. Transformation
3. Validation
4. Featurization

Azure QnA Maker

- Create conversational layer over raw data
- Create knowledge base

Build & Train Model

Regression Model

- Azure Auto ML

Custom Chatbot Model

- Azure OpenAI
- Use GPT-4o as pretrained model
- Set up prompt template
- Answers questions if not found in Knowledge Base
- Validation and Testing

Deploy & Predict

Deploy on Azure OpenAI

Azure Bot Service

Web App

- Next.js
- Azure App Service

Processed data stored back into database for reusability

Time Series Data

Question-answer pairs for finetuning

Use regression model predictions to make recommendations

Search for FAQs and use existing answers

Returns answers found in knowledge base

Interaction with User

