AI Data Story

Use AI to Predict the price of housing in the city of New Orleans.

AI Service

Predict the price of housing based on various characteristics and indicate whether the asking price is above or below the central tendency:

• Continuously monitor changes in the environment around the potential housing stock.

• When new instances of properties for sale are entered, extract the characteristics from the geographic location and the neighborhood.

• Determine if predicted price exceeds the threshold, the central tendency (mean or median).

• Signal to the user that the property is above or below the predicted market price based on available data.

Data Rules

•Encrypt/mask non-public information (NPI) fields before data moves along the workflow

•Encrypt data at rest and transit using SSE and SSL on Azure

•Encrypt data at database level using TDE at SQL Server or Azure SQL Database

•Use all available data to calibrate the model

•Check for important attributes that help in determining the market value of property based on neighborhood and location attributes using Statistical methods

•Build micro-services for automatic information processing, in particular:

• Data cleansing

• Storing

• ETL (Imputing)

• Machine learning pipelines

• Future enhancement for continuous streaming using Azure components

• Reporting using Microsoft PowerBI

•Use Azure Cloud disaster recovery plans to quickly point to database in a different region in case of disaster

•Build a JSON template of Azure resources and security as code, to build the infrastructure quickly in case of failure of data pipeline in one region

•Secure the data in low cost storage classes in different regions to help in fault tolerant and highly available architecture

•NPI information such as SSN, phone number, name, customer address and TIN should not leave or be placed in containers outside USA.

•Return updated housing price by scouring data for new neighborhood amenities and characteristics such as calls for service within 300 meters, restaurants, parks, schools, etc. and using prediction ML algorithm (forecasting).

•Trigger a communication to relevant parties indicating updated values based on changes in neighborhood amenities, especially crimes and flooding.

Data Features Needed

•Location information from Socrata and other external sources.

•Housing information to develop the model from various sources (Zillow, etc).