Describe what we did

FEATURES

Model training UI

Model training takes input sample size

Pre processing, cleaning and formatting

Model Predict UI

Maps Data point visuals with clustering

Data analysis

EDA charts

Histograms

Heatmaps Correlation matrix

ALGORITHIMS

Clustering (Unsupervised)

-We implemented auto clustering when the sample size is more than 5000.

-We implemented Fast marker clusters function on the map

Decision Trees (Supervised)

-We implemented decision trees with decision tree classifier function

Linear Regression (Unsupervised)??

Naïve Bayes (Supervied)??

SVM?

Random forest?

PROCEDURES

CREATING THE MODEL

* We take a random sample
* We preprocess and clean up and format
* We split train and test based on input
* We pick a model
* We fit the model
* We predict the test group to determine the accuracy

PREDICTING

* We input parameters to predict
* We use input parameters and predict using the model
* The prediction is shown

DATA ANALYSIS

* While training, the application also analyses the random number samples
* Histogram, correlation, and EDA is shown

INTERACTIVE DATA VISUALIZATION USING GEOMAPPING

* The data points are plotted in the map using their geo coordinates
* Details of the points are shown in a popup
* Automatic Clustering is implemented using fastmarkercluster function from folium to optimize speed of rendering the map.
* Automatic clustering is set for sample size of greater than 5,000 points