Battle of the Neighborhoods

To build a brand new sports facility in Hyderabad

IBM DataScience Capstone
April, 2019
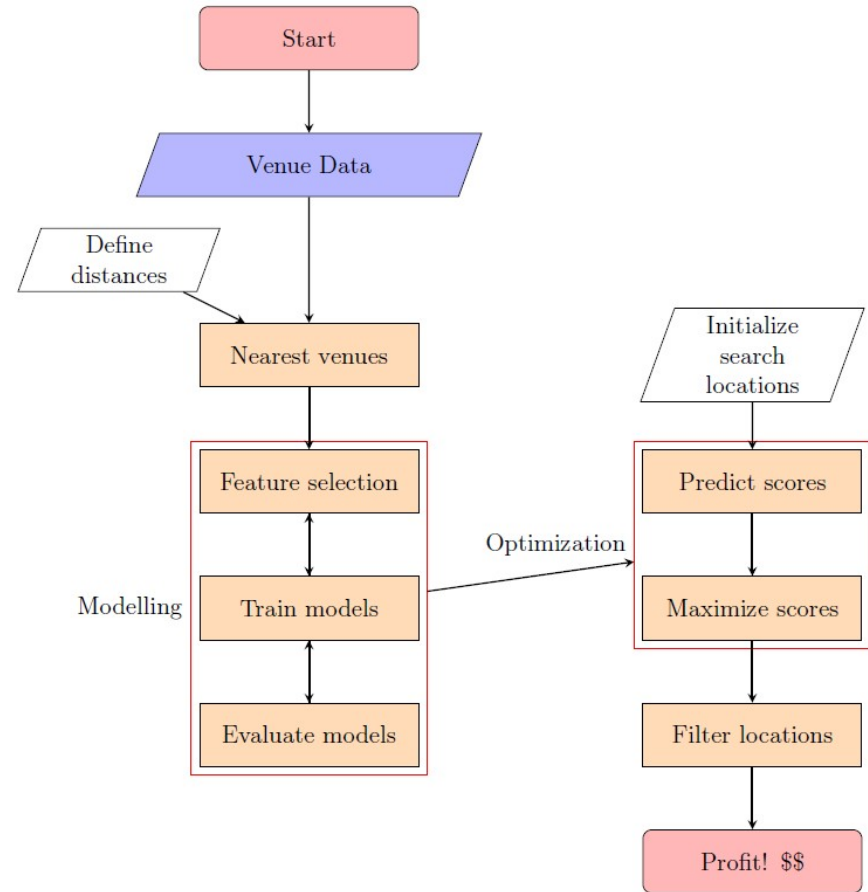Dr. Sabarish Vadarevu

Motivation

- Lack of sports facilities in Hyderabad
- Health implications of sedentary lifestyles
- Transferrable skills from sports competitions
- Stress-relief from engaging in sports

Problem Statement

- Where should a new sports facility be built?
    - Use data on different kinds of venues to characterize neighborhoods
    - Relate neighborhood characteristics to suitability for sports facilities
    - Identify most suitable locations for sports facilities

Methodology

- Collect data using Google Places API

- Train models using Scikit-Learn

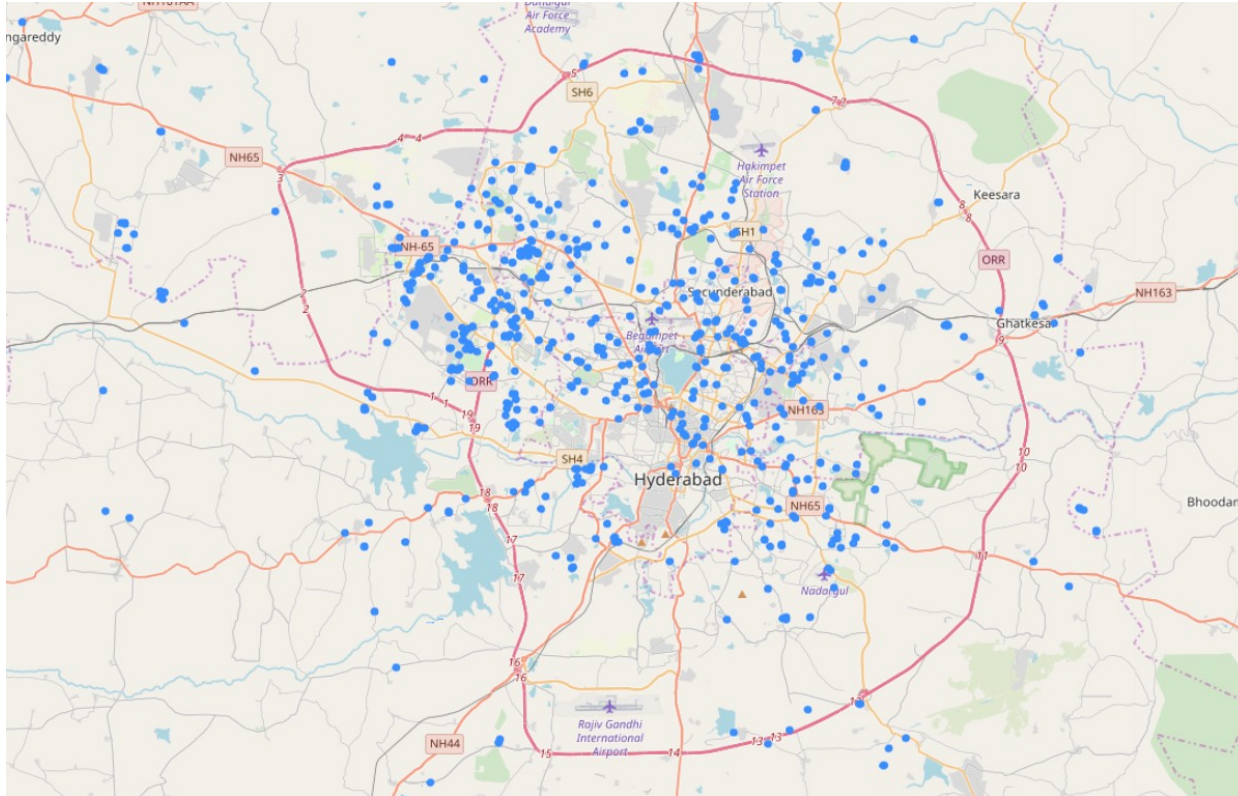- Find optimal locations using Scipy.Optimize

Data Collection

- Data on 10 categories of venues:

Banks	Cinemas
Colleges	Gyms
Hospitals	Pubs
Restaurants	Schools
Sports Facilities	Supermarkets

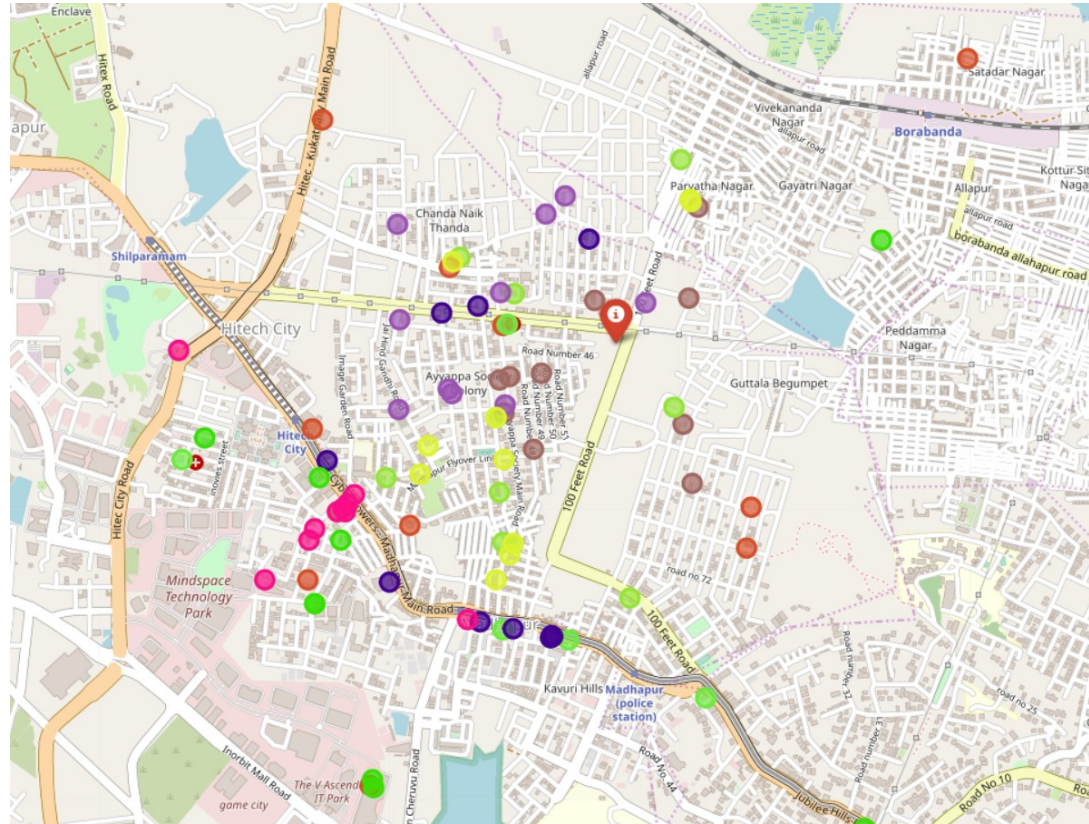# Data Sample (Banks)

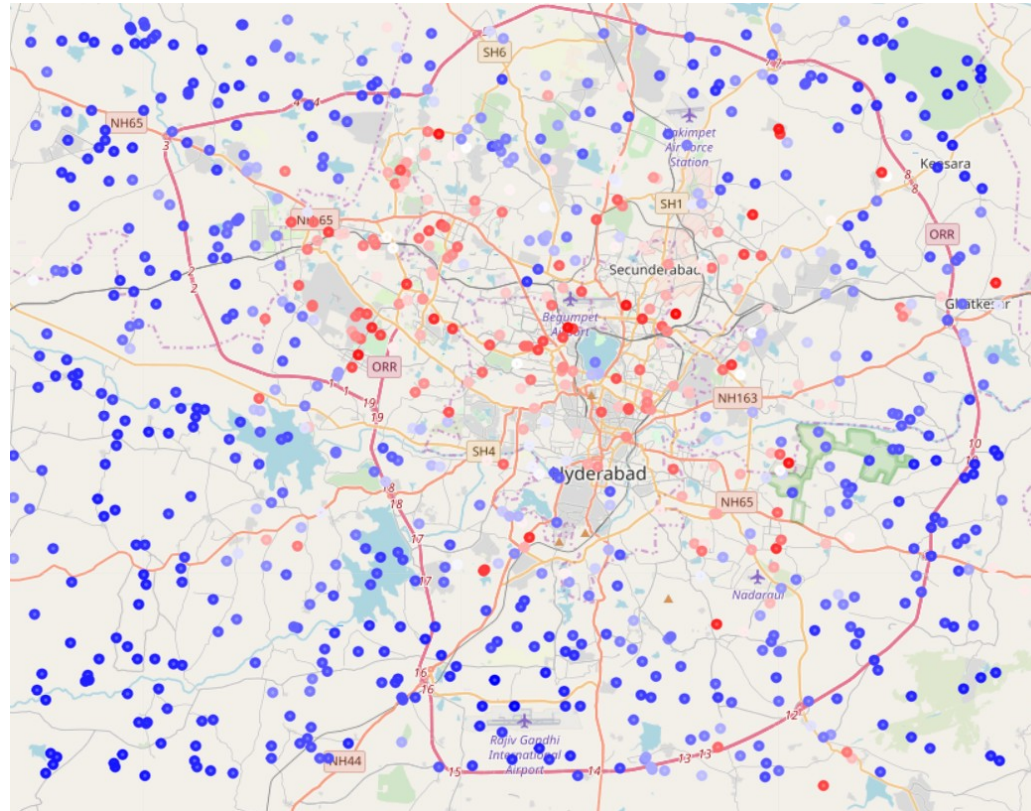| Name	ID	Latitude	Longitude	Dist_CC
The Hongkong An	ChIJvwWaYkyXyzsRbwJxN4y4bso	17.418570	78.461631	1390
Union Bank Of I	ChIJL_wg6X6XyzsRr3PTAhKPr_0	17.423035	78.462451	1514
Vijaya bank Cor	ChIJQaJ7btyXyzsRcgjjyVtbZ2g	17.418486	78.457346	1832
IDBI Bank	ChIJXba01s6ZyzsRFAXQFyNR7jA	17.390077	78.492433	3404
State Bank E-Co	ChIJJ43Qi4miyzsRpGuqwmnZLdw	17.319765	78.518138	11568

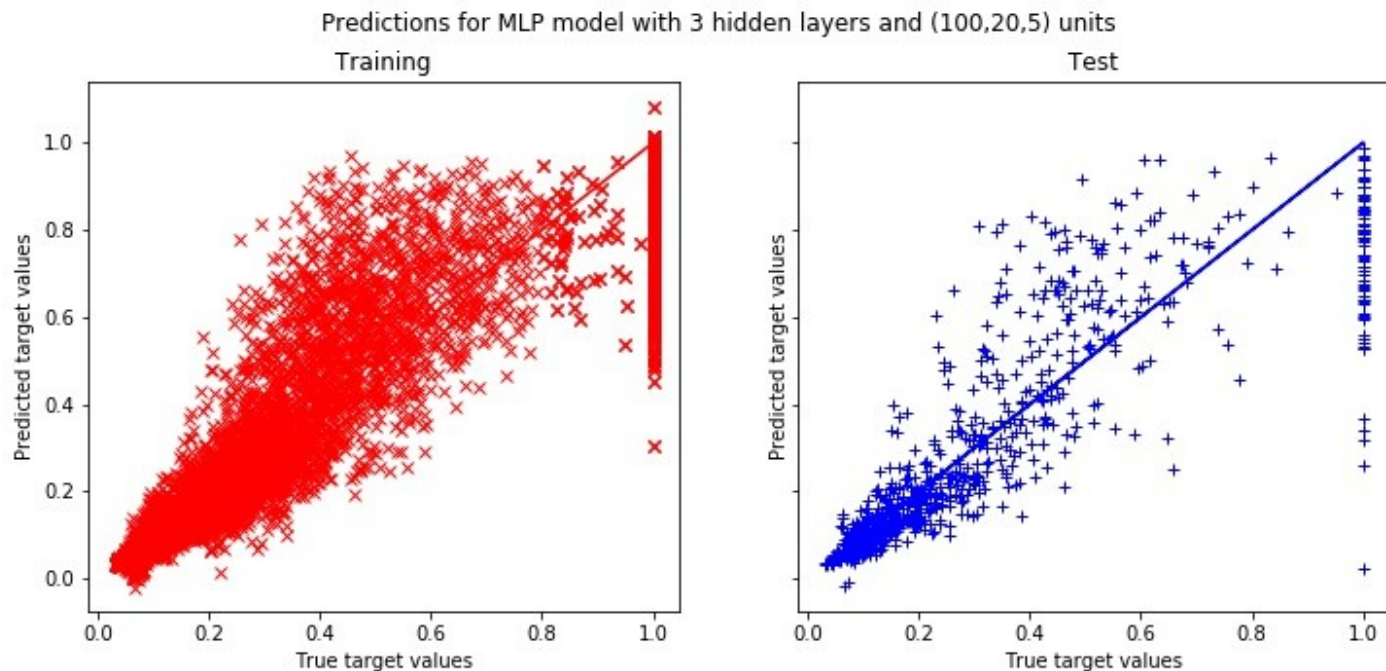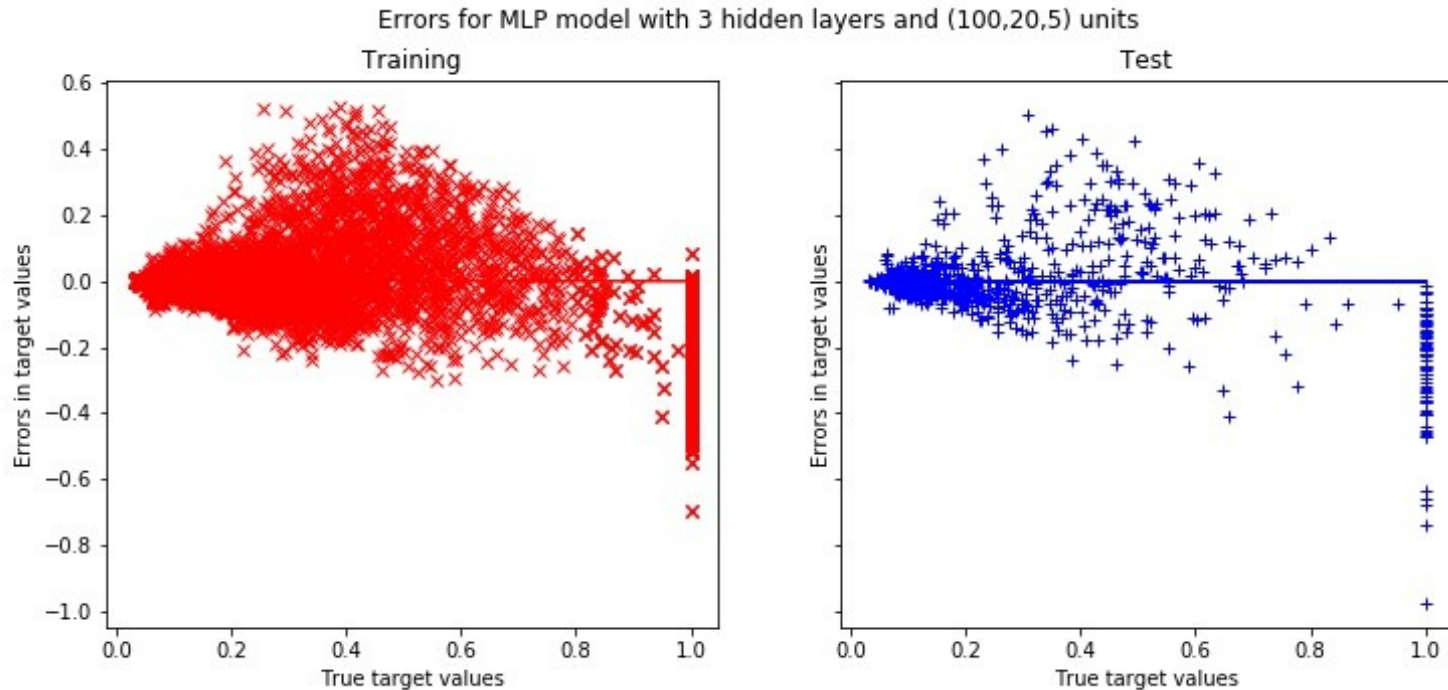# Data Sample (Sports Facilities)
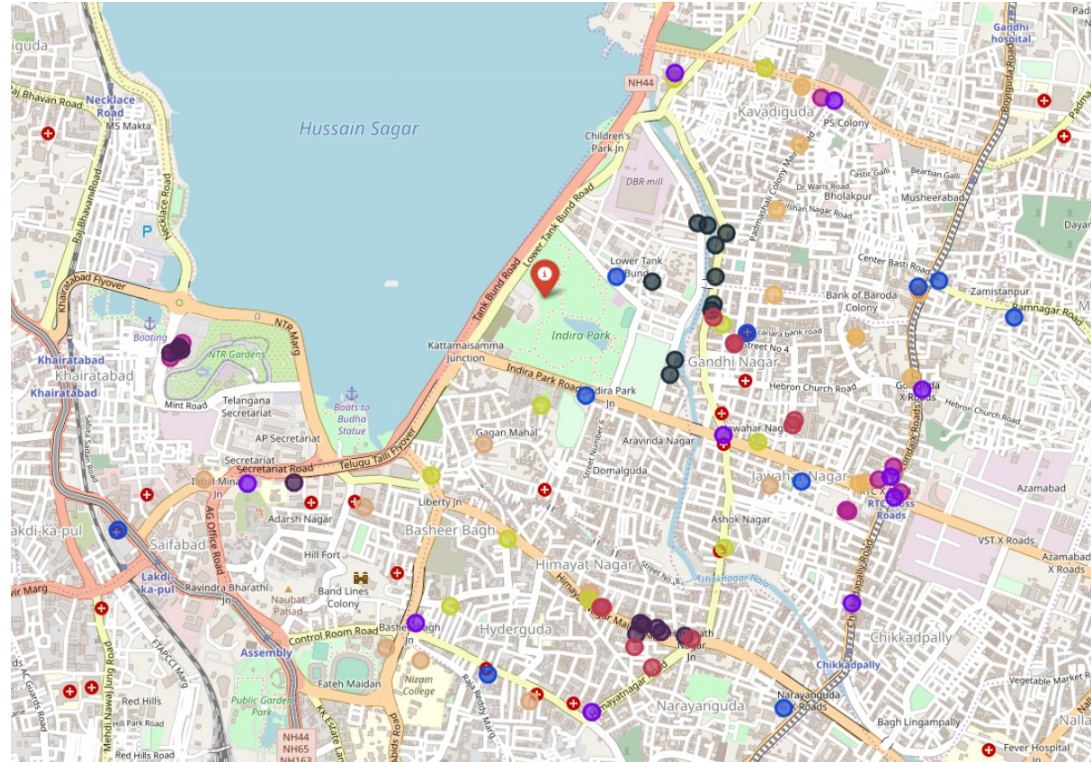
Nearest Venues

Suitability Scores

Best Model – Multilayer Perceptron

Predictions for MLP model with 3 hidden layers and (100,20,5) units

Best Model – Multilayer Perceptron

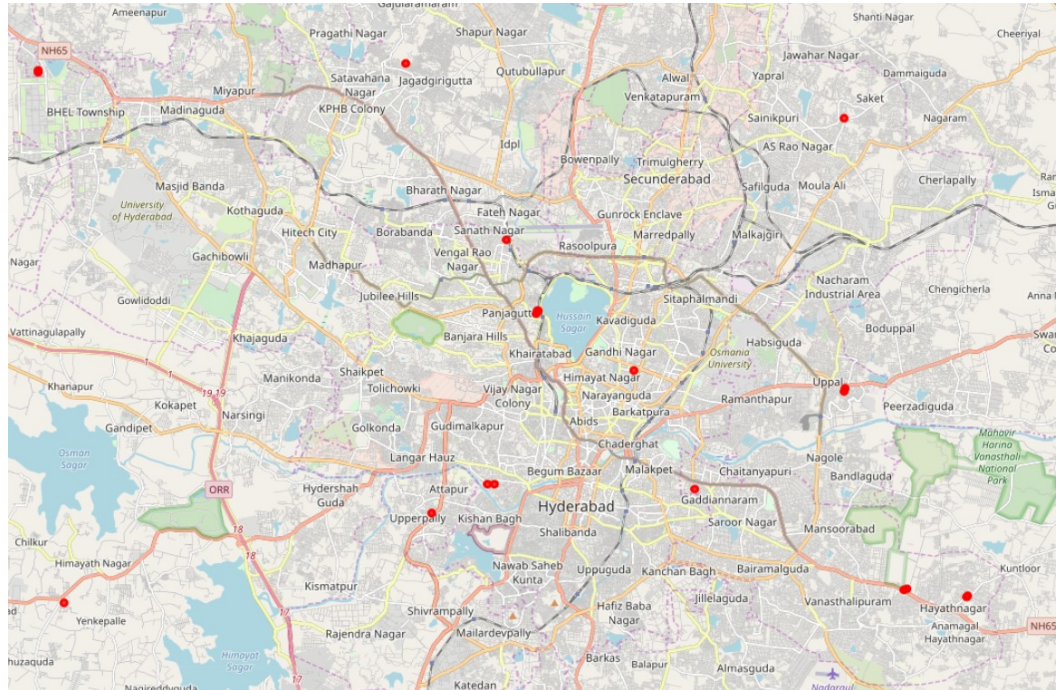Errors for MLP model with 3 hidden layers and (100,20,5) units

Optimal Location (near City Center)

Optimal Locations (Best Twenty)

Thank You