

Hursh Karkhanis

Data Analyst <> Machine Learning <> Data Visualization

www.hurshkarkhanis.com

SUMMARY

Data analyst with project experience in data analysis, statistical testing, data visualization, and machine learning skills, specifically regression, classification, and neural networks. Fluent in SQL and Python.

DATA ANALYSIS AND MACHINE LEARNING PROJECTS

Linear Regression and Feature Importance: NBA2K21 Ratings

- Cleaned, split, and fit player and rating data to Linear Regression model with low RMSE
- Ranked feature importance to see which statistical categories had the most impact on 2K rating

Sentiment Analysis and Statistical Tests: 2020 Election Tweets

- Found spike in sentiment for both candidates after Trump COVID diagnosis
- Won first prize in Galvanize Annual Datathon

Classification and LSTM: Classifying Tweets by Sentiment

- Used TFIDF vectorizer to vectorize tweet text for classification models
- Trained sk-learn classification models using an AWS EC2 instance, highest accuracy score was 0.8

Classification and Profit Curves: Fraud Detection Project

- Trained classification models with numerical and dummified categorical data
- Built profit curves estimating the dollar gain/loss of true/false predictions

Linear Regression: Predicting IMDb Ratings

- Engineered features based on domain knowledge
- Conducted GridSearch to improve Gradient Boosting model. Final RMSE was half of baseline

Statistics: "Defund the Police" statistical tests

- Found statistically significant relationship between police spending and crime rate
- Rejected null hypothesis based on low p-value

EDUCATION

University of San Francisco

2016-2020

Major: *Bachelors of Science Business Administration, Finance* with a minor in *Computer Science*

Galvanize Data Science

2020

Exploratory Data Analysis, Statistical Tests, Regression and Classification Machine Learning