Predicting Low Back Pain Severity

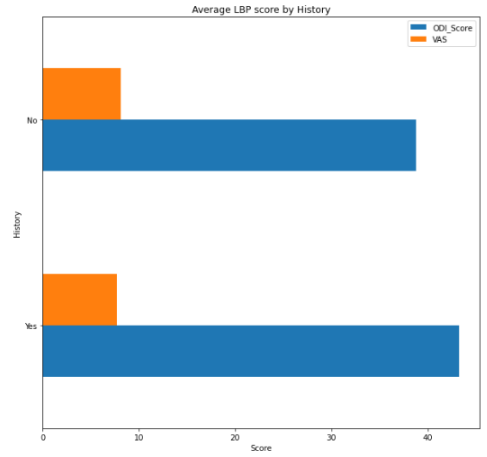
Capstone Project

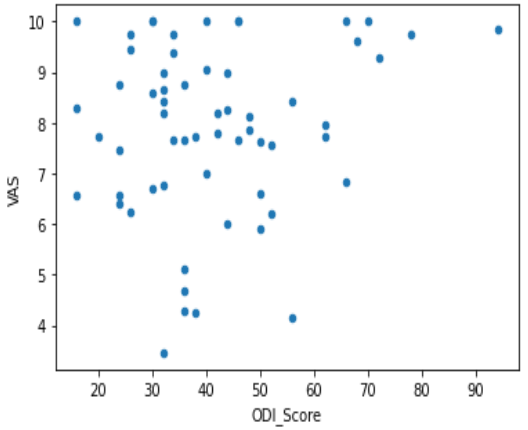
Introduction

- LBP is one of the most common disorders
- Three primary risk factor categories for LBP:
  - Occupational
  - Personal
  - Psychosocial
- Predictive model can help in developing effective intervention methods by determining severity

Exploratory Data Analysis

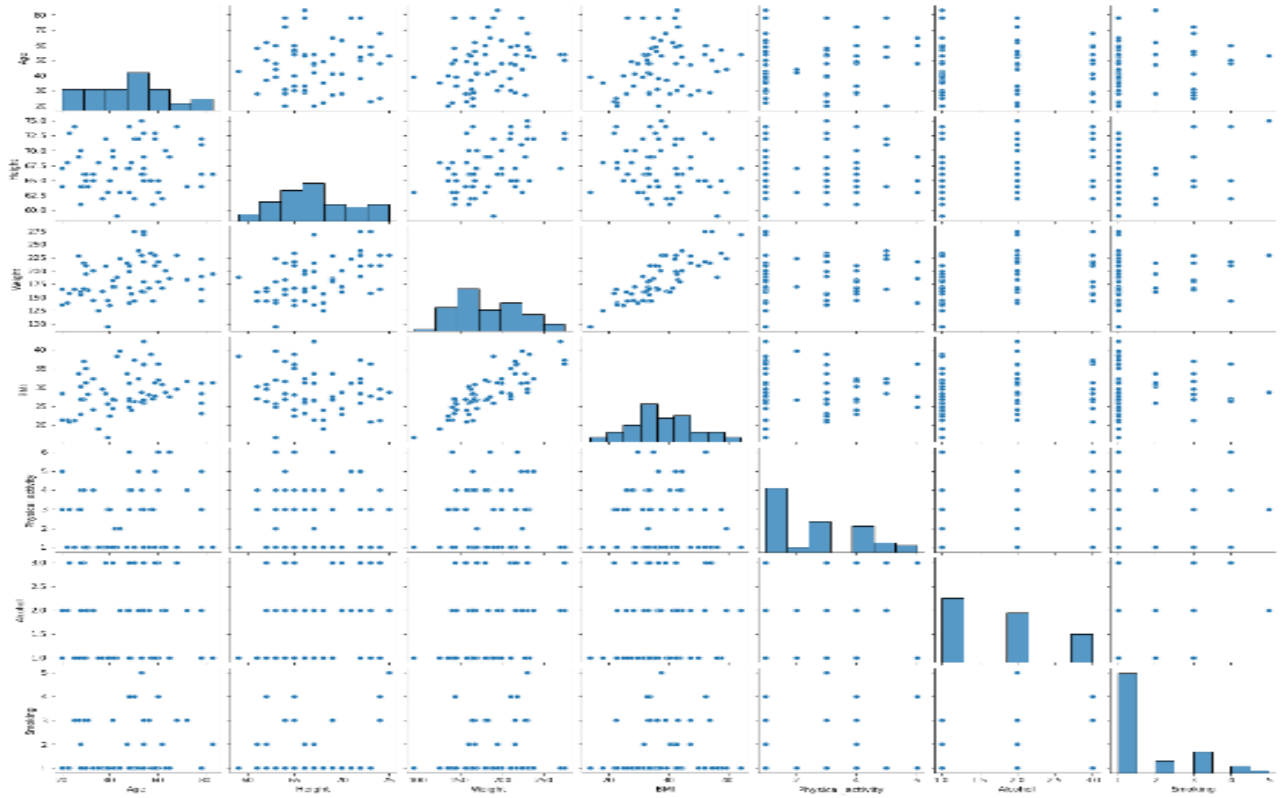
- 60 rows and 28 columns
- 2 measures of LBP severity – VAS and ODI Score
- ODI chosen for model building





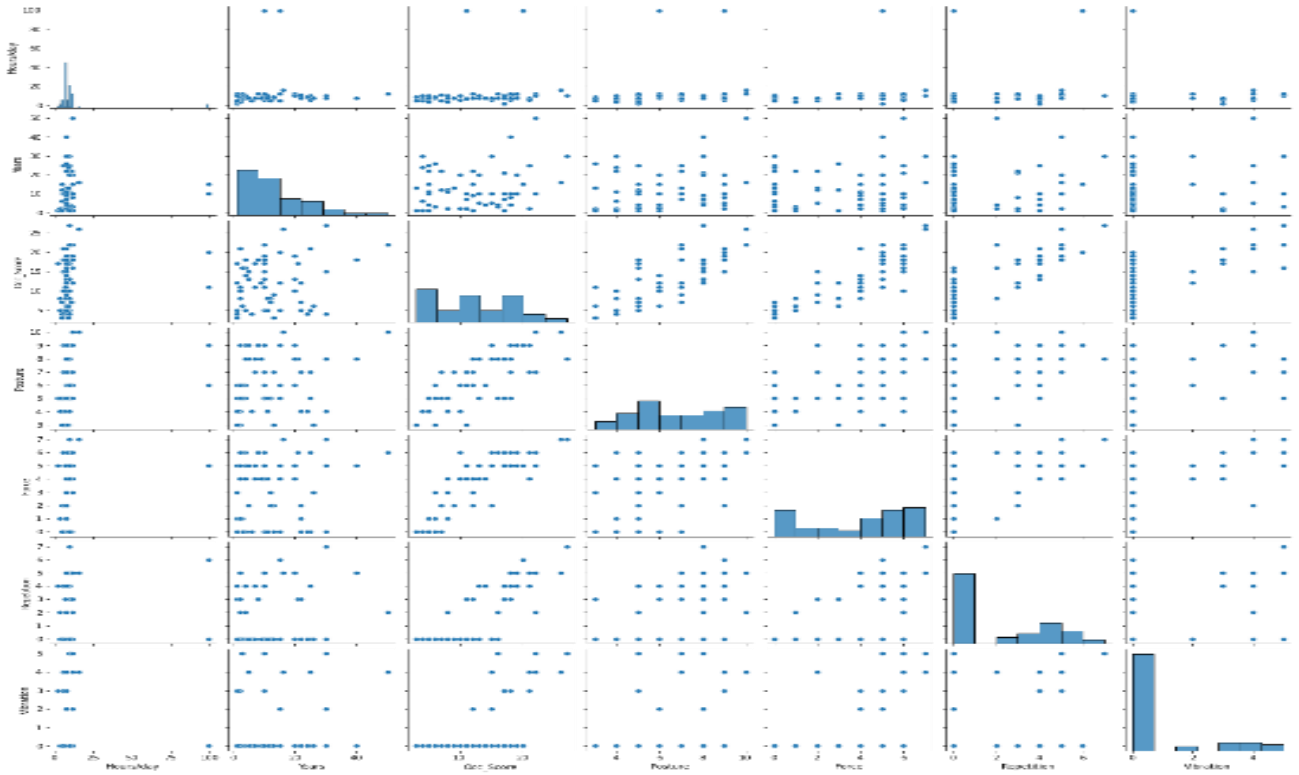
Exploratory Data Analysis

- Correlations between Personal Factors



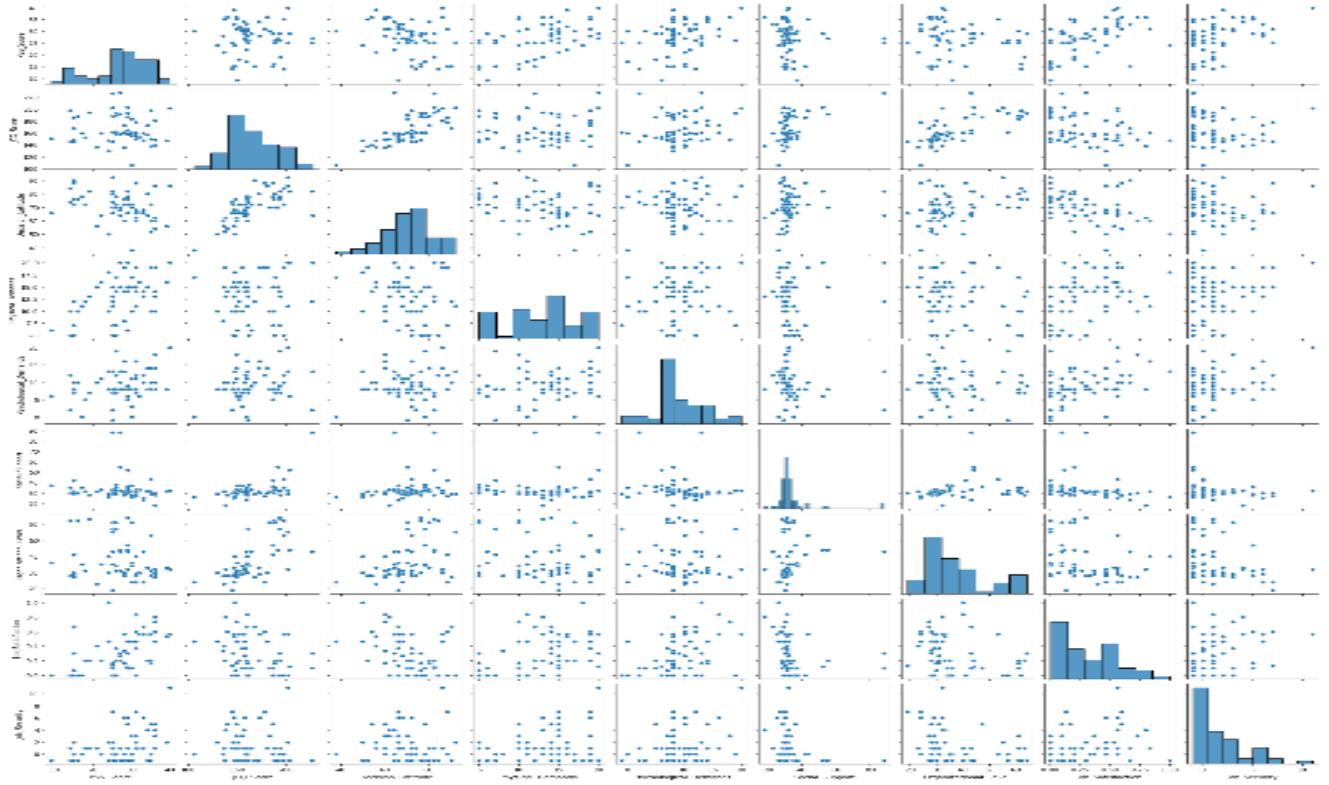
Exploratory Data Analysis

- Correlations between Occupational Factors



Exploratory Data Analysis

- Correlations between Psychosocial Factors



Feature Selection & Preprocessing

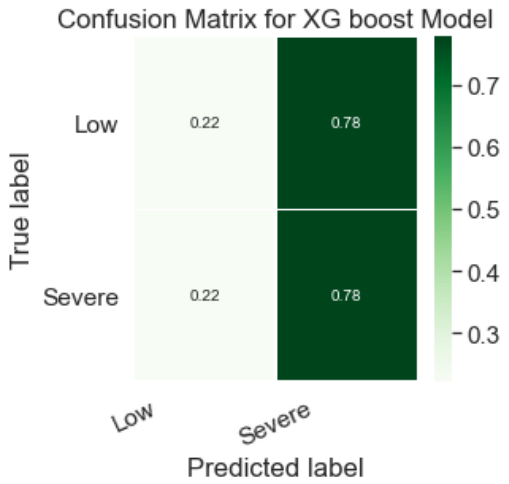
- Missing values filled with median values
- Features were converted into categorical variables
- Binary model built with ODI2 – split into 0 and 1 for low and high LBP severity
- Tertiary model built with ODI – split into 0, 1 and 2 for low, moderate and high LBP severity
- One hot encoding on categorical variables - Gender, History, Ethnicity, BMI, Physical activity etc

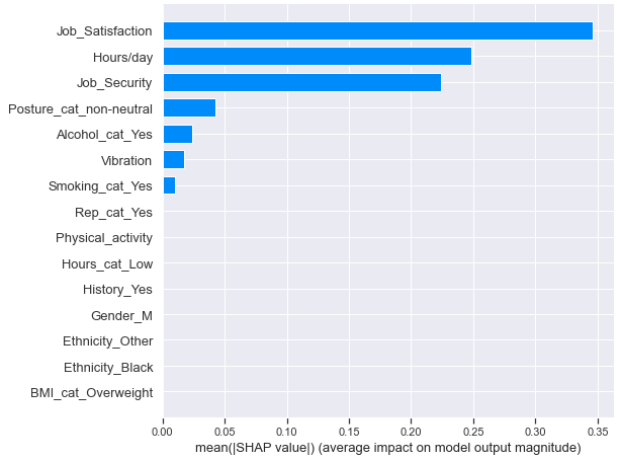
Modelling

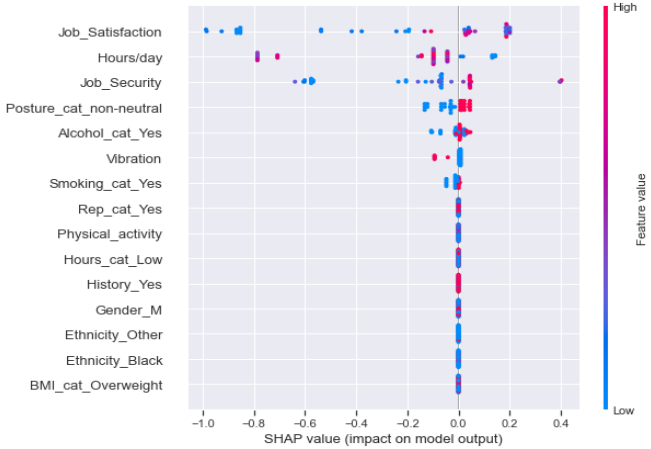
- Data split into 70/30 training and testing sets
- XGBoost (eXtreme Gradient Boosting) used for classification models
- Random Forest Regression used as regression model

Results

- Binary Model – accuracy score of 0.5

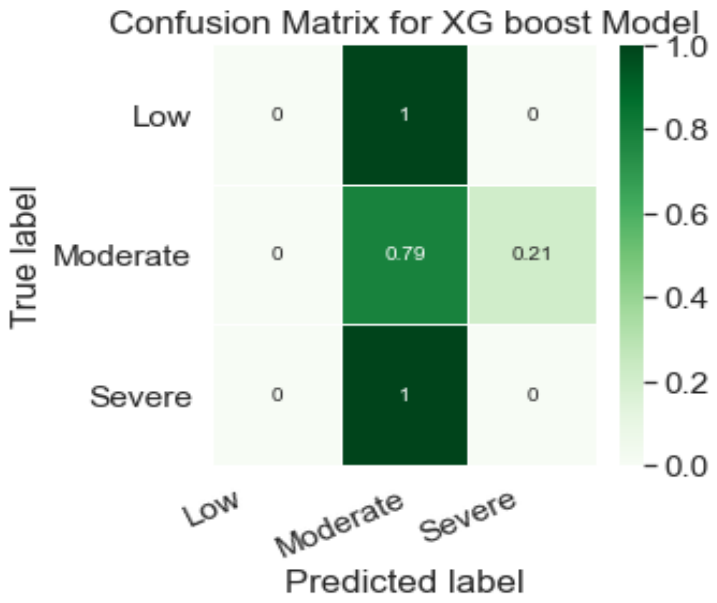


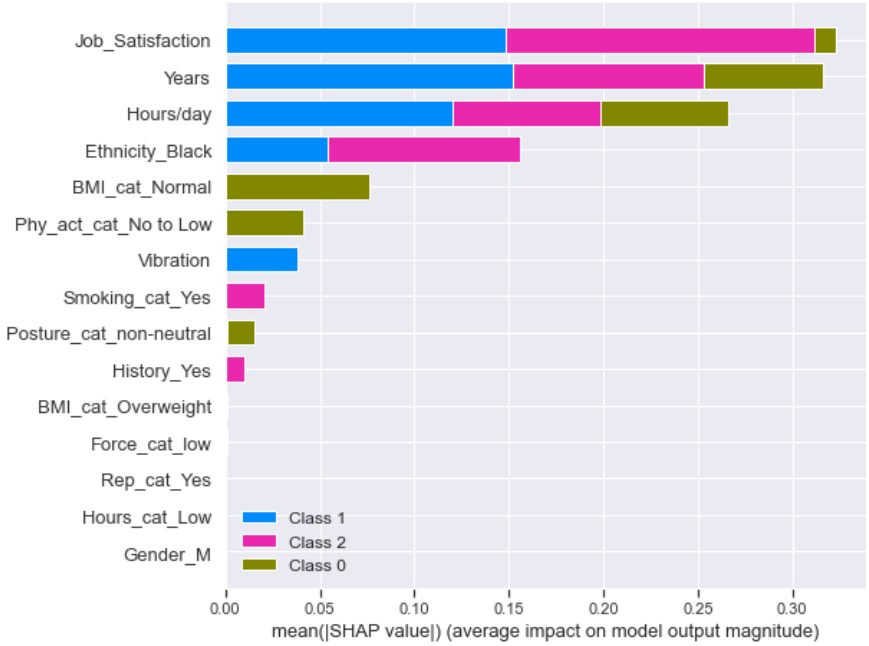




Results

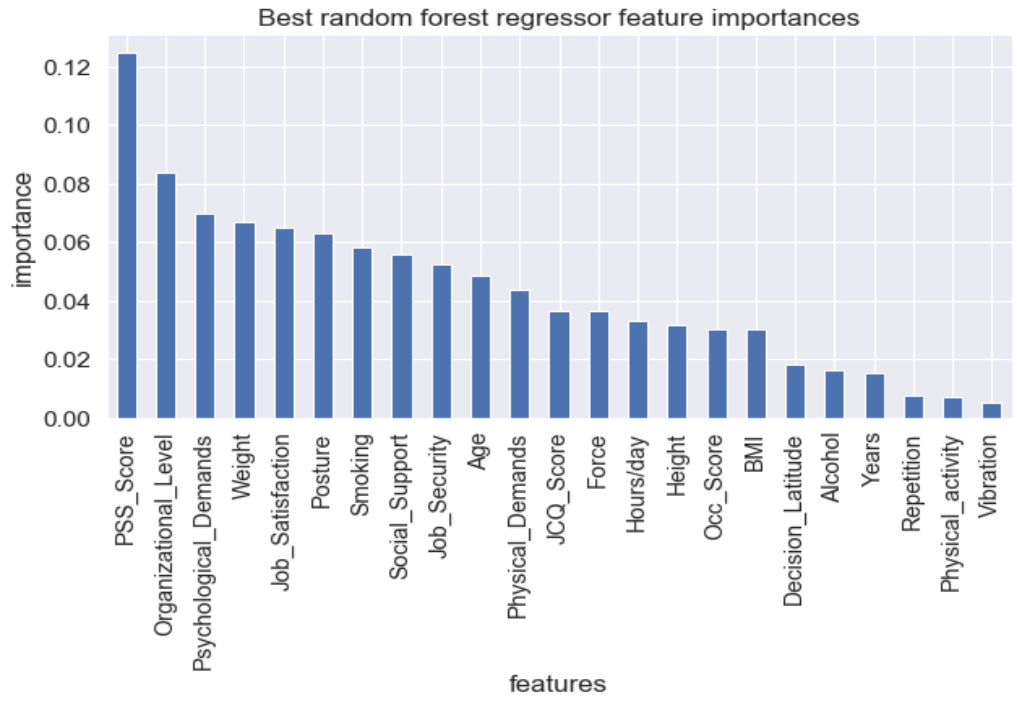
- Tertiary Model – accuracy score of 0.61





Results

- Regression Model –RMSE value of 5.6



Discussion & Conclusions

- Tertiary classification - best predictive model
- Job Satisfaction with negative correlation - important indicator in both classification models
- PSS - most important in the regression model
- Psychosocial factors need to be considered when designing preventive measures
- One limitation - limited sample size
- More efficient model could be probably developed using a larger sample