Python for Data Analysis

Landslide after Rainfall

Victor QIAN
Martin RAMPONT

01

Subject introduction

Landslide after Rainfall - An Understudied Field

Landslide after Rainfall

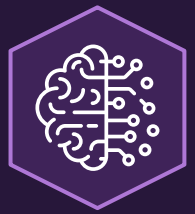
Thousands deaths
Billions $ damages

/ Each year

Use cases of Landslide Analysis

Landslide Suceptibility

Identify areas that are prone to landslides
+ map in real time



Early warning system

Provide timely warnings to residents and authorities.

Impact Assessment

Assess the extent of damage and plan for recovery
(Computer Vision)

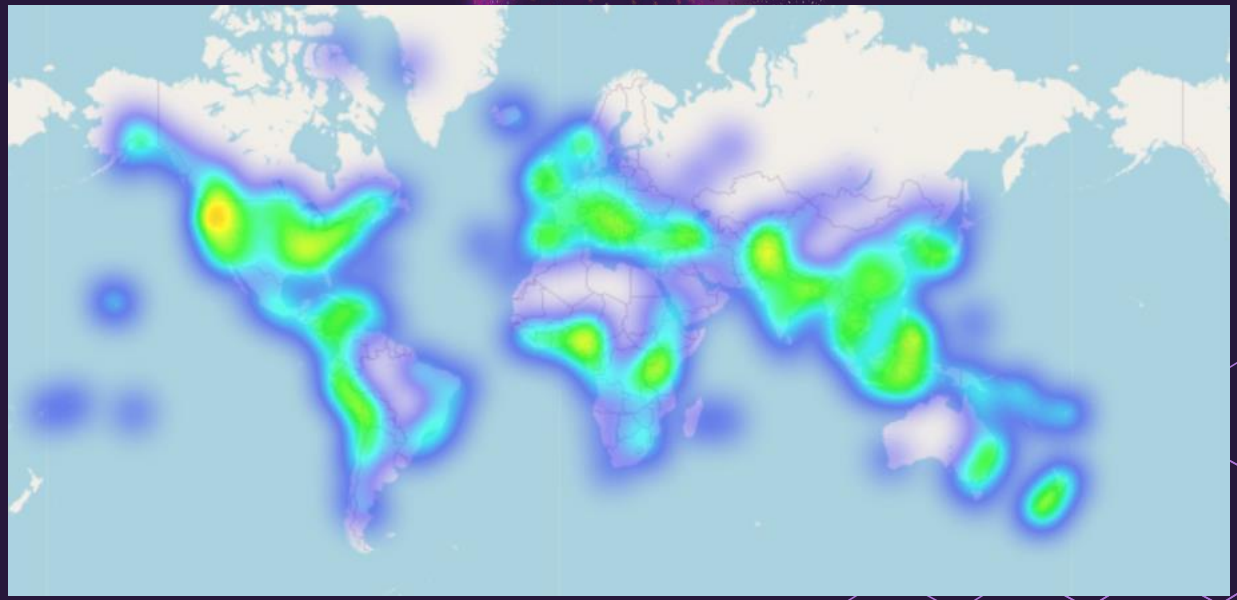
02

The Dataset

Global Landslide Catalog Export (2007-2016), NASA

Global Landslide Catalog Export (2007-2016), NASA

Rows	Columns	Each row is a
11K	31	Landslide

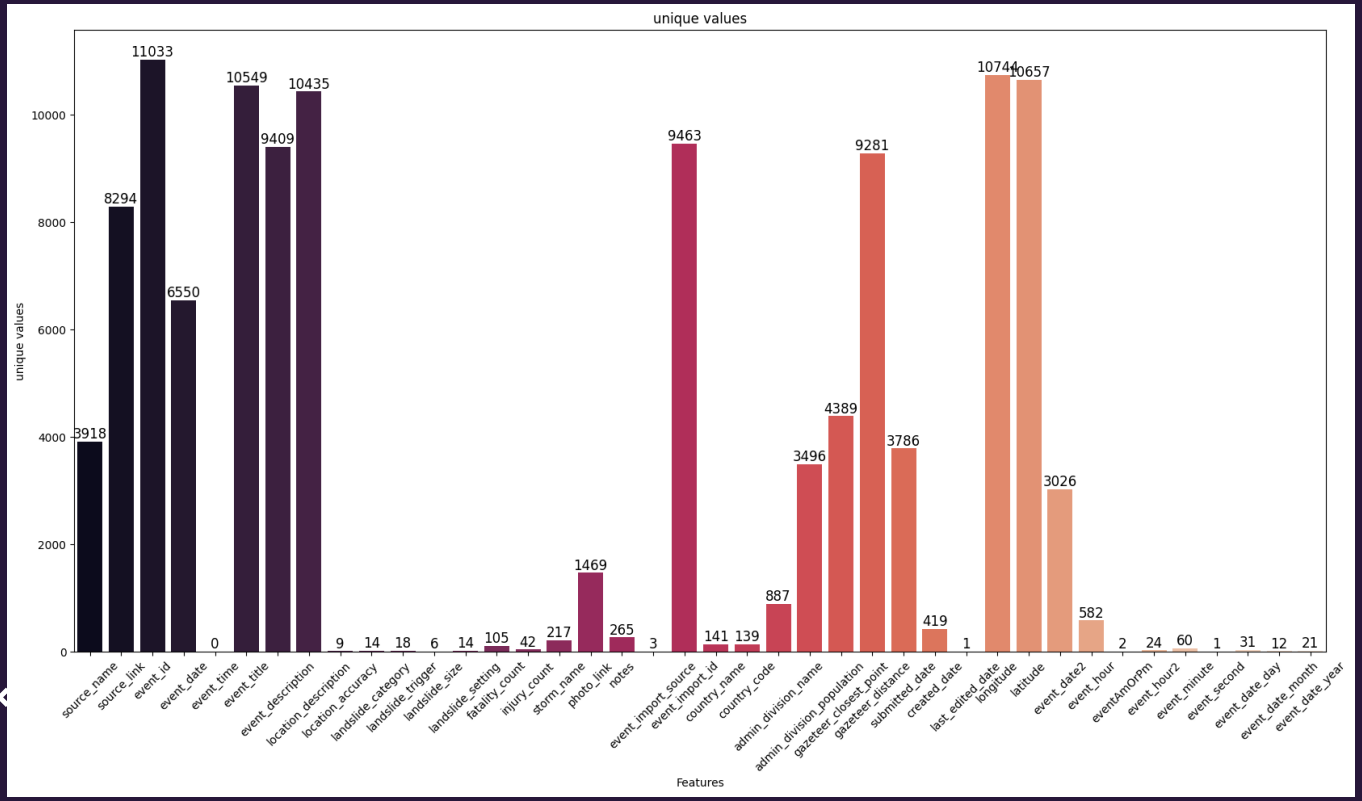


03

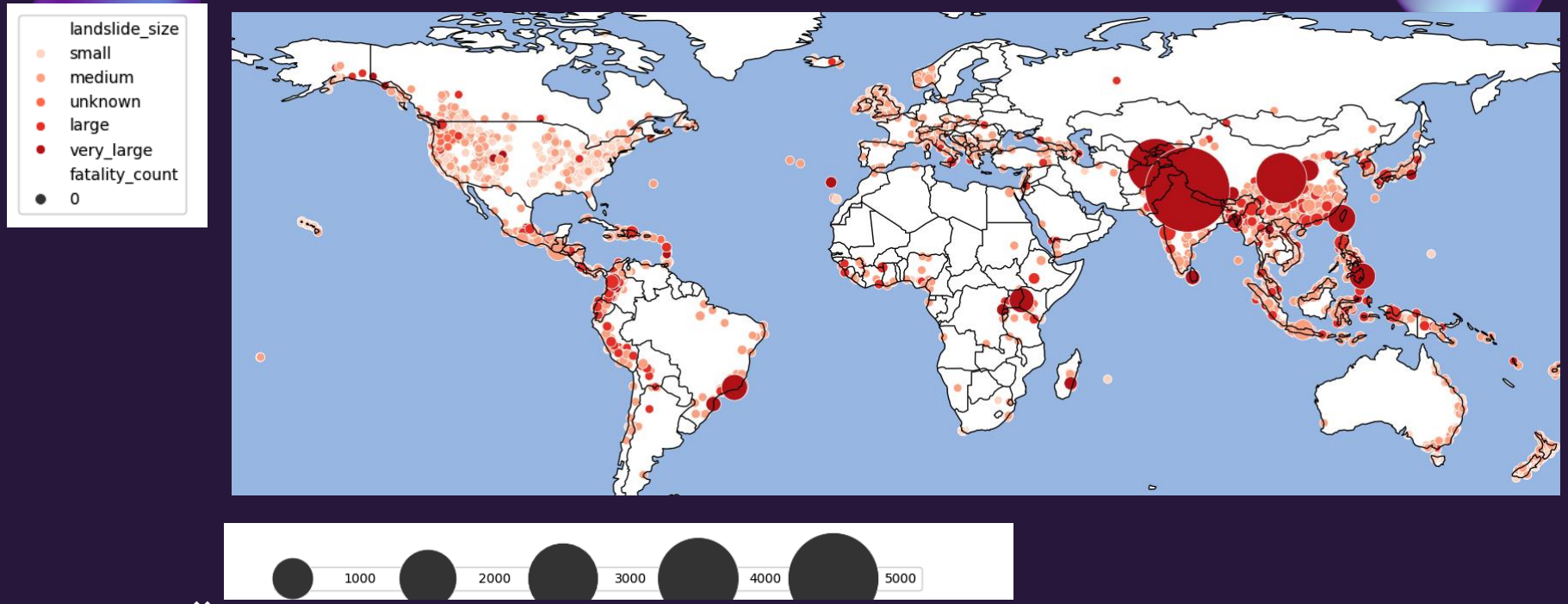
EDA

ASL Recognition

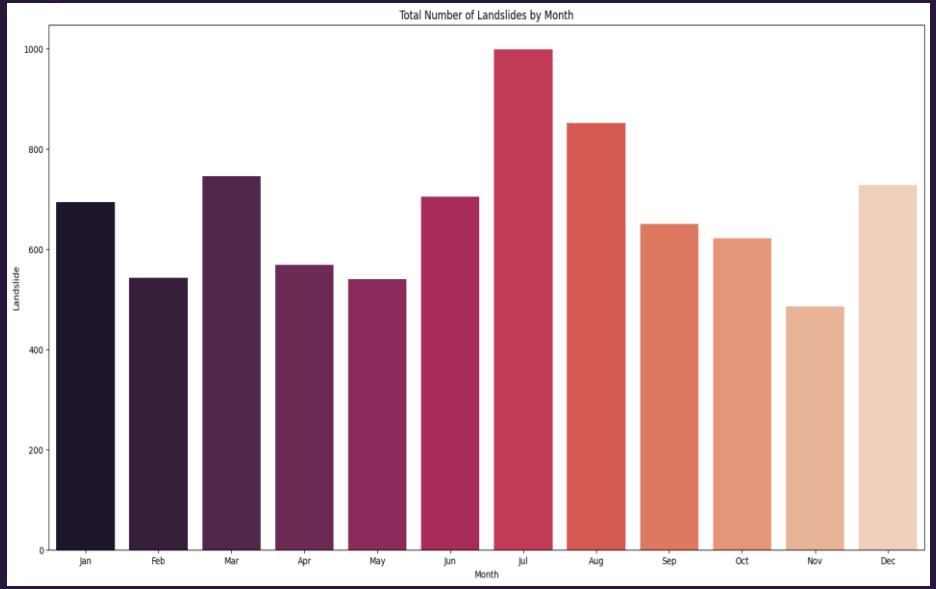
Missing values



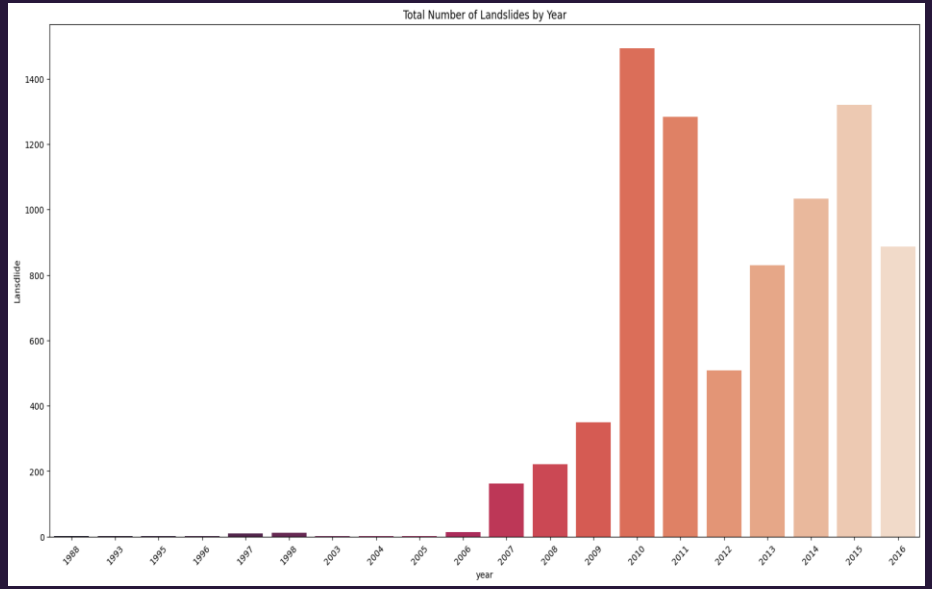
Map of Fatality Count



Time Seasonality

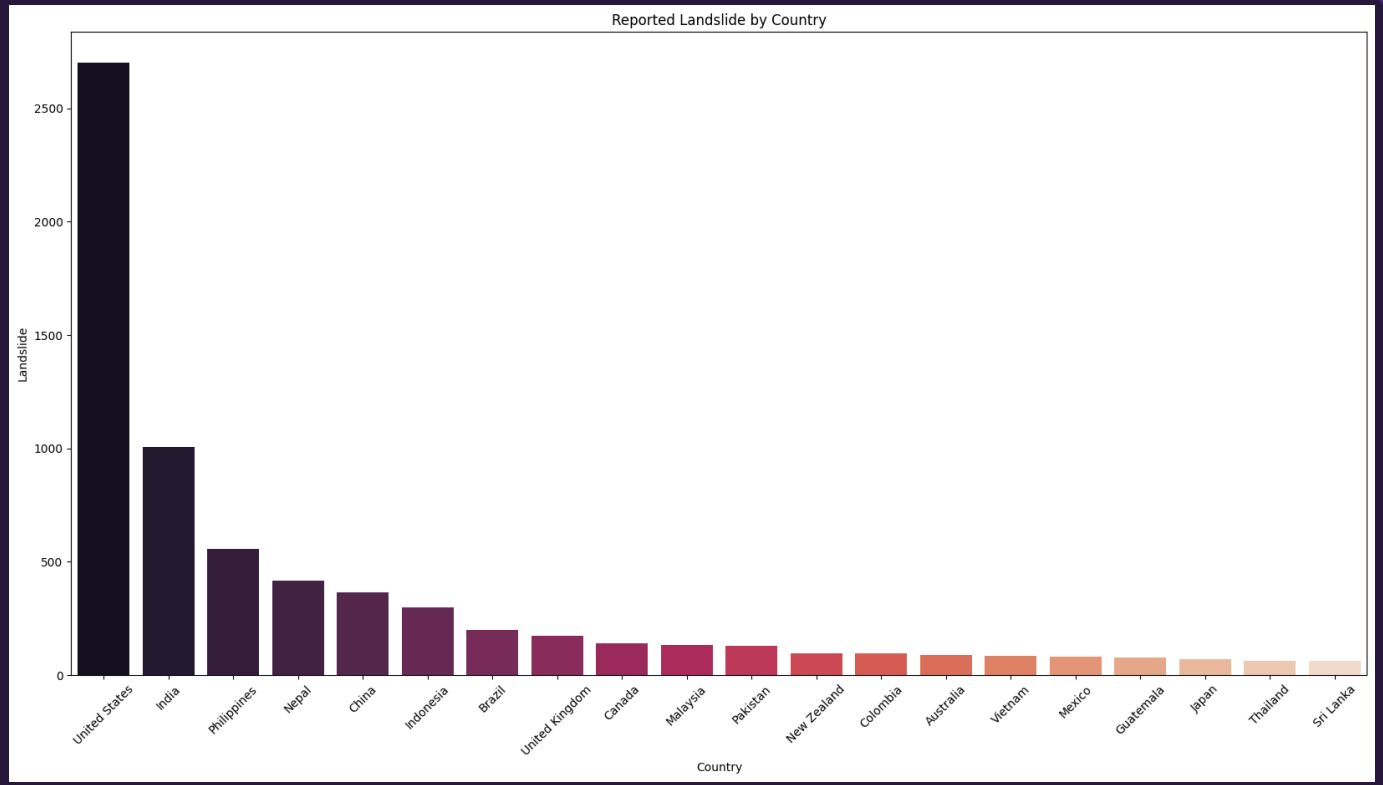


By Month

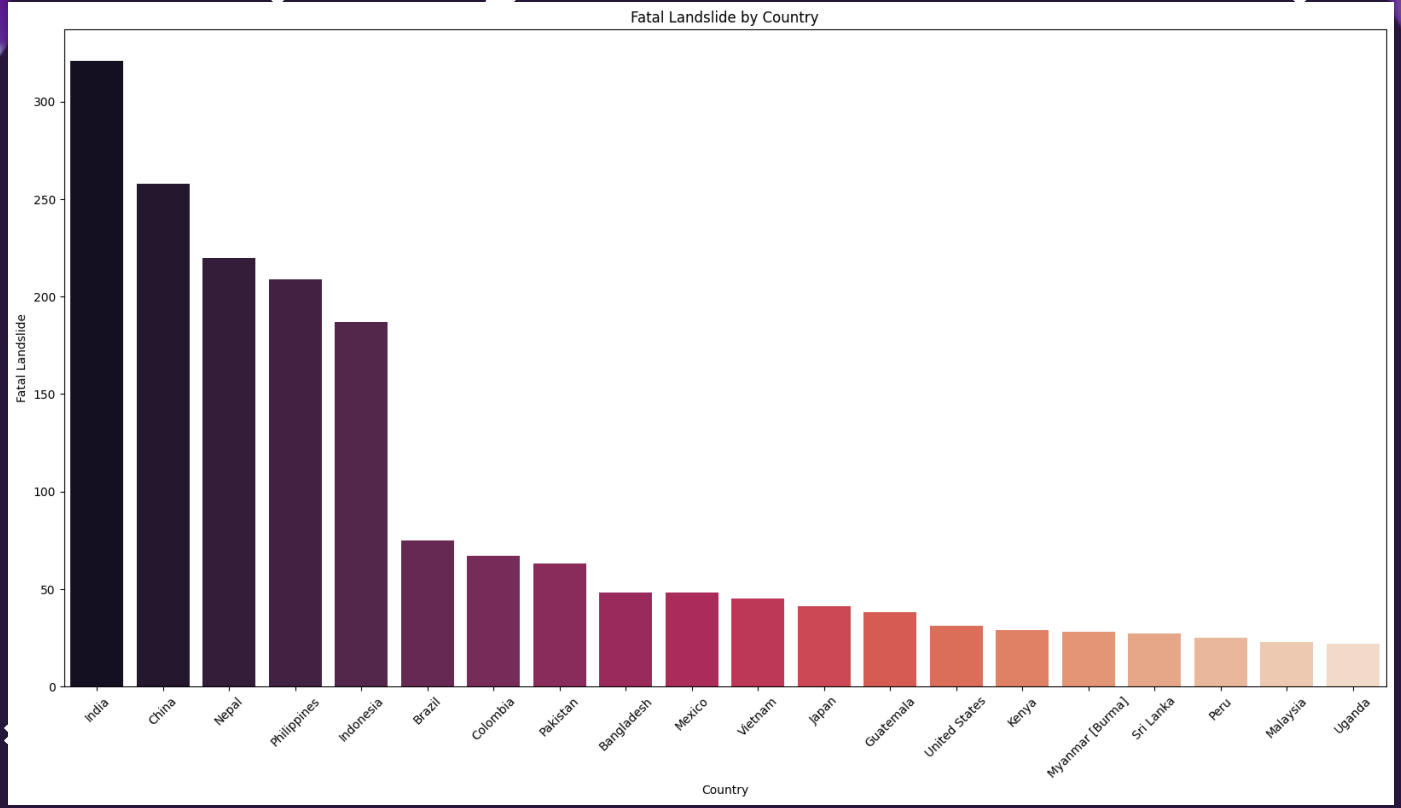


By Year

Country Landslide Number



Country deaths number



04

Machine Learning

ASL Sign Language Recognition

Data Preprocessing

Missing Values

Deals with missing values

Incorrect data types and values

Values Formating

Columns Splitting

Time data splitting

Final output

Normalization

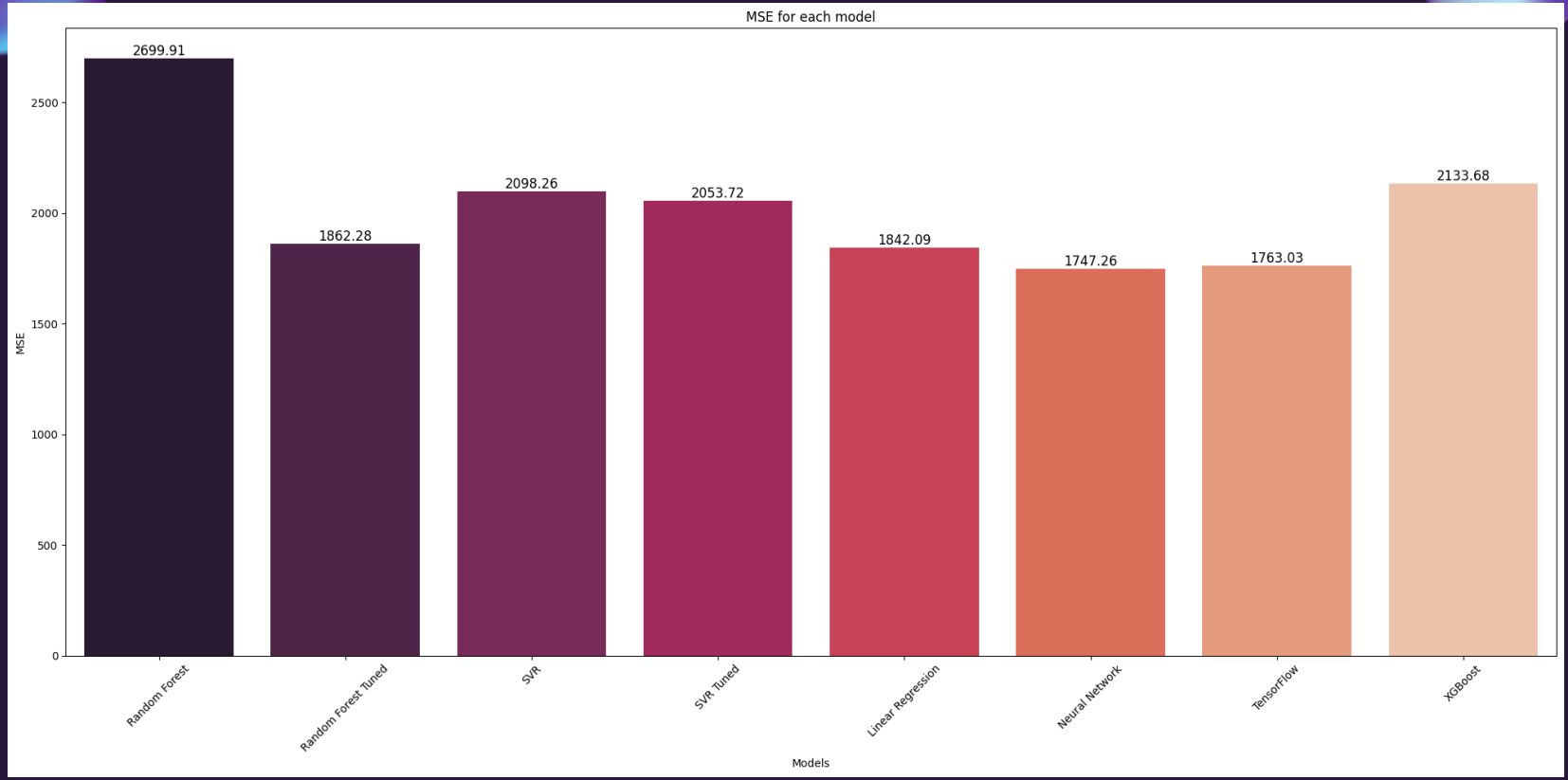
Types of trained models

AI model	Advantages	Drawbacks
Random Forest	Handles non-linearity and interactions well	Can be prone to overfitting, especially on noisy data
Random Forest Tuned (Grid Search)	Improved performance through hyperparameter tuning	Increased computational cost due to tuning
SVR	Effective in high-dimensional spaces	Sensitive to noise and may require careful preprocessing
SVR Tuned	Better generalization with optimized parameters	Computationally expensive, especially with grid search
Linear Regression	Simple and interpretable	Assumes a linear relationship, may not capture complex patterns
Neural Network	Powerful for complex, non-linear relationships	Requires large amounts of data and computation resources
Tensorflow	Comprehensive deep learning framework	Steeper learning curve, especially for beginners
Xgboost	Handles missing values and outliers well	Prone to overfitting, may require tuning

05 Results

ML Results

Types of trained models



THANKS!

DO YOU HAVE ANY QUESTIONS ?