Introduction to Machine Learnining

Day 4: Data Preprocessing & Regression

Aditya Mohan
Hands on Mentor
GitHub @(7avenged)

Abhimanyu Aman
IEEE Member

Akhilesh Kumar
Hands on Mentor
GitHub @(akhilesh-k)

Agenda

Day 4

Recap of Day 3

Python Libraries revisted

Data Preprocessing with Scikit Learn

Supervised Learning

Regression

Basic Overview and Hands on Regression

GitHub Resources

Quick Recap

NumPy

Pandas

Matplotlib

Data preprocessing

Numpy #recap

```
import numpy as np
```

np.array([rank])

np.zeros((row x col))

np.ones((row x col))

np.full((row x col), 'constant')

np.eye(size)

np.sum()

np.subtract()

np.multiply()

np.divide()

np.sqrt()

np.dot()

Pandas #recap

```
import pandas as pd

pd.read_csv('file.csv')

pd.to_csv('file.csv')

pd.read_excel('file.xls')

pd.to_excel('file.xls')

.shape

.head()

.describe()

.iloc()
```

Matplotlib #recap

```
import matplotlib.pyplot as plt

plt.plot(x, y, ' ')

plt.xlabel(' ')

plt.ylabel(' ')

plt.title(' ')

plt.show()

plt.subplot( )

fig.savefig()

plt.hist()

plt.scatter()
```

Data Preprocessing #recap

Importing and Reading Datasets

Statistical processing

Data Encoding

Feature scaling

Manipulating Datasets

Machine Learning

Classes of ML algorithms:

Supervised Learning

Unsupervised Learning

Reinforcement Learning

Deep Learning

What we will cover

Regression

- Linear Regression: Simple and Multiple
- Polynomial Regression
- Decision Tree
- Random Forest

Classification

- Logistic Regression
- KNN
- Support Vector Machines
- Naive Bayes
- Decision Tree

Thank You for attending

This session was presented by ACM JUIT