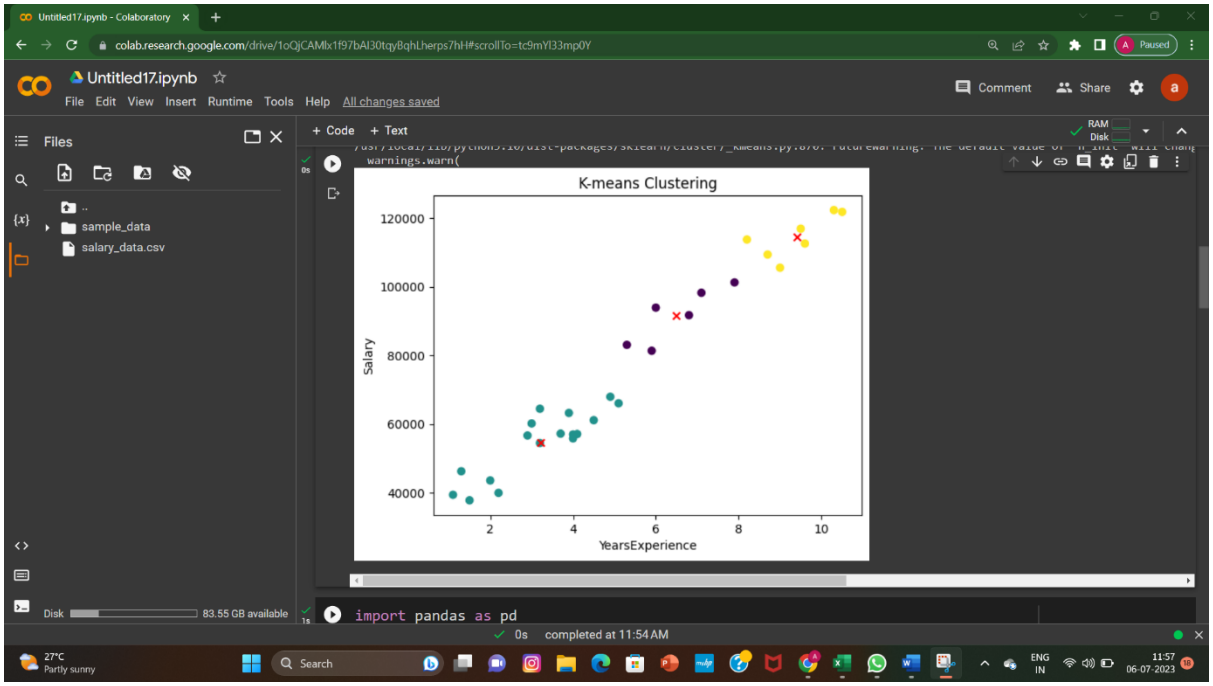
225 Vaibhav Jadhav

239 Atharv Lokhande

Lab assignment 6

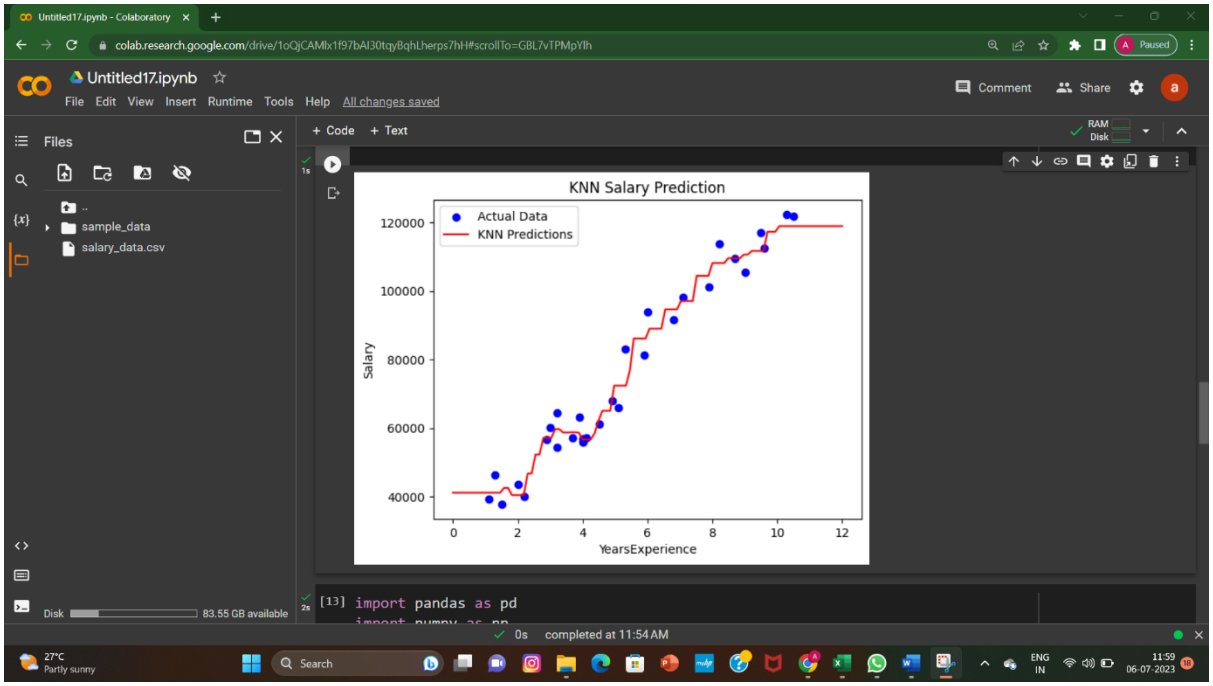
k-means clustering.

```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

# Read the dataset from a CSV file
df = pd.read_csv("/content/salary_data.csv")

# Extract the relevant columns for clustering
X = df[['YearsExperience', 'Salary']].values

# Perform k-means clustering
k = 3  # Number of clusters
kmeans = KMeans(n_clusters=k)
kmeans.fit(X)

# Get the cluster labels and centroids
labels = kmeans.labels_
centroids = kmeans.cluster_centers_

# Visualize the clusters
plt.scatter(X[:, 0], X[:, 1], c=labels, cmap='viridis')
plt.scatter(centroids[:, 0], centroids[:, 1],
marker='x', color='r')
plt.xlabel('YearsExperience')
plt.ylabel('Salary')
plt.title('K-means Clustering')
plt.show()
```

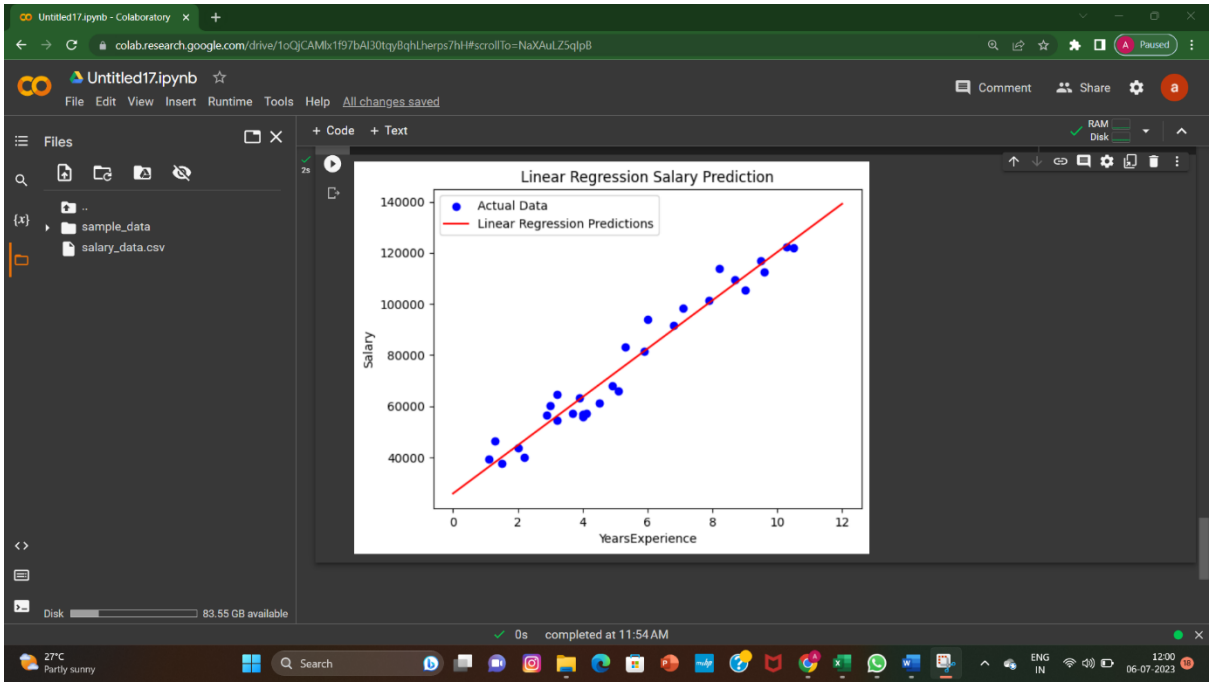

Untitled17.ipynb - Colaboratory
colab.research.google.com/drive/1oQjCAMlx1f97bAI30tqyBqhLherps7hH#scrollTo=tc9mYI33mp0Y
Untitled17.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
salary_data.csv
Disk 83.55 GB available
+ Code + Text
RAM
Disk
warnings.warn(
K-means Clustering
Salary
YearsExperience
120000
100000
80000
60000
40000
2
4
6
8
10
import pandas as pd
0s completed at 11:54 AM
27°C
Partly sunny
Search
ENG
IN
11:57
06-07-2023

#KNN

```
import pandas as pd
import numpy as np
from sklearn.neighbors import KNeighborsRegressor
import matplotlib.pyplot as plt

# Read the dataset from a CSV file
df = pd.read_csv("/content/salary_data.csv")

# Extract the relevant columns for prediction
X = df[['YearsExperience']].values
y = df['Salary'].values

# Perform KNN prediction
k = 3  # Number of nearest neighbors
knn = KNeighborsRegressor(n_neighbors=k)
knn.fit(X, y)

# Generate salary predictions for the given years of
experience
years_of_exp = np.linspace(0, 12, 100).reshape(-1, 1)
salary_predictions = knn.predict(years_of_exp)

# Visualize the KNN predictions
plt.scatter(X, y, color='b', label='Actual Data')
plt.plot(years_of_exp, salary_predictions, color='r',
label='KNN Predictions')
plt.xlabel('YearsExperience')
plt.ylabel('Salary')
plt.title('KNN Salary Prediction')
plt.legend()
plt.show()
```


Untitled17.ipynb - Colaboratory
colab.research.google.com/drive/1oQjCAMlx1f97bAI30tqyBqhLherps7hH#scrollTo=GBL7vTPMpYIh
Untitled17.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
salary_data.csv
Disk 83.55 GB available
+ Code + Text
RAM
Disk
KNN Salary Prediction
Actual Data
KNN Predictions
Salary
YearsExperience
120000
100000
80000
60000
40000
0
2
4
6
8
10
12
[13] import pandas as pd
0s completed at 11:54 AM
27°C
Partly sunny
Search
ENG
IN
11:59
06-07-2023

```
import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt

# Read the dataset from a CSV file
df = pd.read_csv("/content/salary_data.csv")

# Extract the relevant columns for prediction
X = df[['YearsExperience']].values
y = df['Salary'].values

# Perform linear regression
regression = LinearRegression()
regression.fit(X, y)

# Generate salary predictions for the given years of
experience
years_of_exp = np.linspace(0, 12, 100).reshape(-1, 1)
salary_predictions = regression.predict(years_of_exp)

# Visualize the linear regression predictions
plt.scatter(X, y, color='b', label='Actual Data')
plt.plot(years_of_exp, salary_predictions, color='r',
label='Linear Regression Predictions')
plt.xlabel('YearsExperience')
plt.ylabel('Salary')
plt.title('Linear Regression Salary Prediction')
plt.legend()
plt.show()
```


Untitled17.ipynb - Colaboratory
colab.research.google.com/drive/1oQjCAMlx1f97bAI30tqyBqhLherps7hH#scrollTo=NaXAuLZ5qIpB
Untitled17.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
salary_data.csv
+ Code + Text
RAM
Disk
Linear Regression Salary Prediction
Actual Data
Linear Regression Predictions
Salary
YearsExperience
140000
120000
100000
80000
60000
40000
0
2
4
6
8
10
12
Disk 83.55 GB available
0s completed at 11:54 AM
27°C
Partly sunny
Search
ENG
IN
12:00
06-07-2023