



[71]:

import

numpy

as

np

import

pandas

as

pd

from

sklearn.model_selection

import

train_test_split

from

sklearn.feature_extraction.text

import

TfidfVectorizer

from

sklearn.linear_model

import

LogisticRegression

from

sklearn.metrics

import

accuracy_score

[72]:

data=pd.read_csv('Downloads//salary_of_employee.csv')

data.head()

[72]:

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
0	32.0	Male	Bachelor's	Software Engineer	5.0	90000.0
1	28.0	Female	Master's	Data Analyst	3.0	65000.0
2	45.0	Male	PhD	Senior Manager	15.0	150000.0
3	36.0	Female	Bachelor's	Sales Associate	7.0	60000.0
4	52.0	Male	Master's	Director	20.0	200000.0

[73]:

data.tail()

[73]:

	Age	Gender	Education Level	Job Title	Years of Experience	Salary
370	35.0	Female	Bachelor's	Senior Marketing Analyst	8.0	85000.0
371	43.0	Male	Master's	Director of Operations	19.0	170000.0
372	29.0	Female	Bachelor's	Junior Project Manager	2.0	40000.0
373	34.0	Male	Bachelor's	Senior Operations Coordinator	7.0	90000.0
374	44.0	Female	PhD	Senior Business Analyst	15.0	150000.0

[74]:

data.shape

[74]:

(375, 6)

[75]:

data.dtypes