



JOB Salary Prediction IN CSE Field

Group - 4

Our team



**MD Faiz
Nawaz**

ID : 2013964642



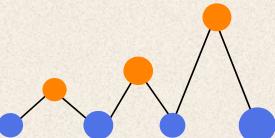
**Md. Khairul
Alam Munna**

ID: 2022484642



Md Al Amin

ID: 1811904042



Project General Description

- Our project focuses on developing a machine learning model for predicting salaries
- The objective is to utilize various data features to accurately estimate the expected salary for a given job position.
- This involves collecting and preprocessing data, selecting appropriate machine learning algorithms, training the model, and evaluating its performances



Hypothesis

01

Employees in the technology sector might have higher salaries compared to other sectors.

02

Employees with more experience tend to have higher salaries.

03

Job satisfaction might correlate with the likelihood of considering job switching.

Feature Description

01

Joining Date

Calendar

02

Job Title

Software Developer

03

Required Experience

3 Category

04

Salary

Options

05

Industry Sector

Technology

06

Demanding of Job

Increasing

Feature Description

07

Overall Satisfaction

Satisfaction

08

Working Model

Onsite

09

Job Switching

Yes/No

10

Reasons

Better Salary

Continuation of Dataset

01

Cleaning the Data

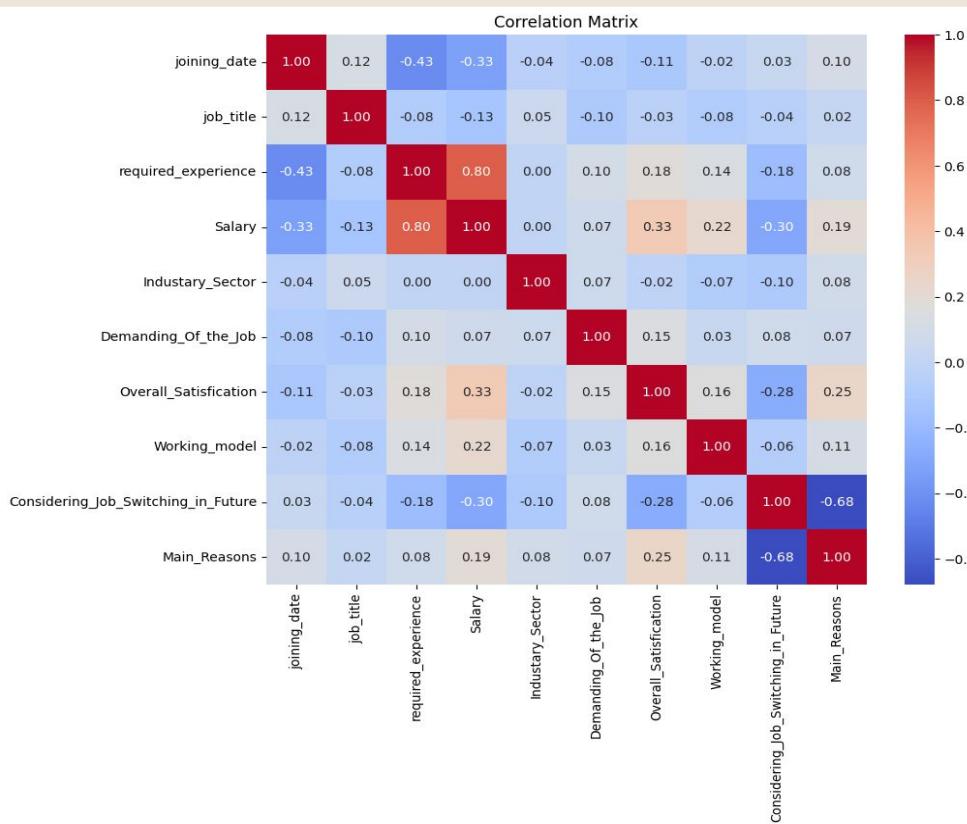
02

Normalization

03

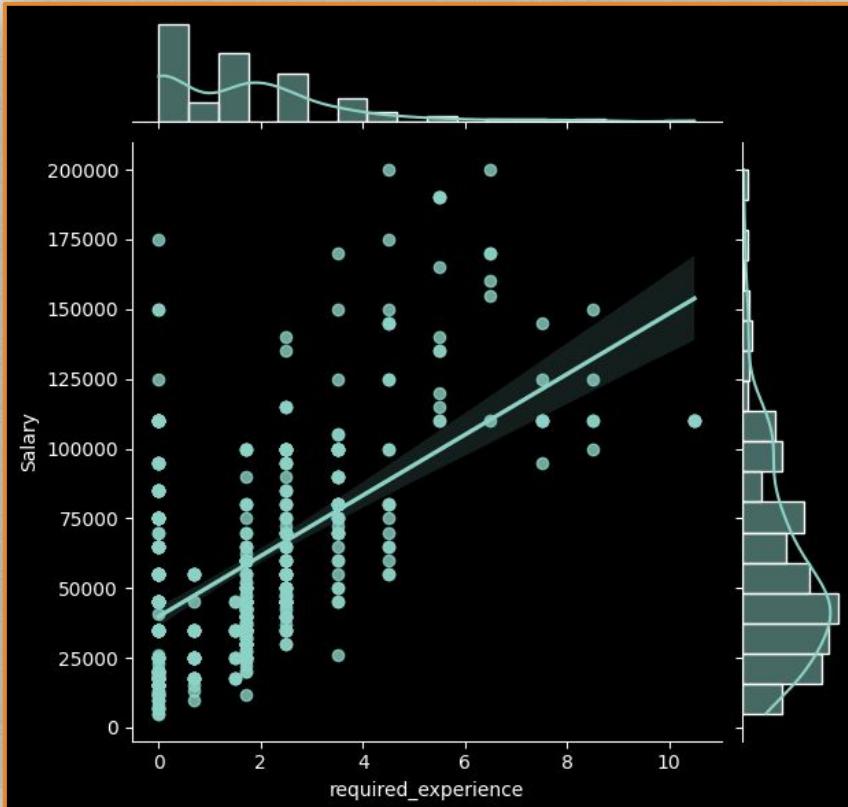
Outlier Removal

Correlation Analysis



Target Function
Salary

Joint Plot Analysis



Positive Correlation

As experience rises, so does salary.

Feature Selections

Forward

Incremental addition enhances model accuracy.

Backward

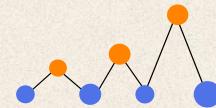
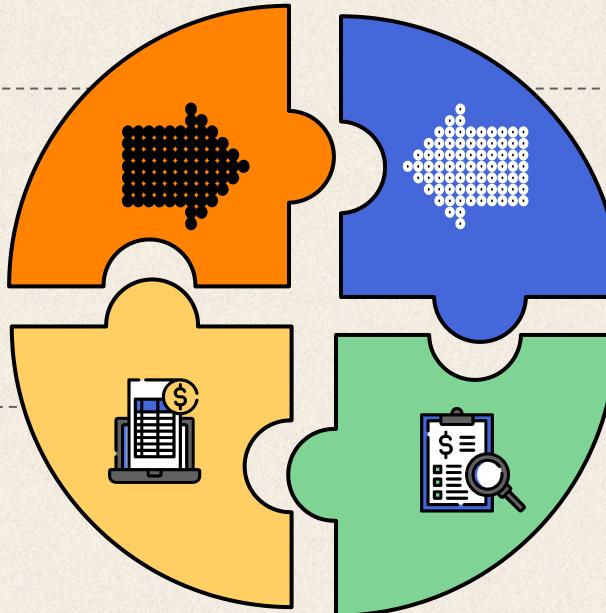
Systematic removal refines model precision.

Correlation Matrix

Strong correlations improve predictive strength

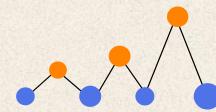
Hypothetical

Feature combinations enhance model interpretability.



Result Presentation

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree	39.5 & 65	62.4	65.6	63
KNN	60	57	55.8	54.8
Support Vector Machine	40 & 54.4	37	28	29
Linear Regression	50.14	Nan	Nan	Nan
Logistic Regression	63	64	58	63





Train/ Test/ Validation

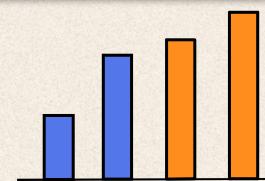


```
1 from sklearn.model_selection import train_test_split
2 from sklearn.tree import DecisionTreeClassifier
3 from sklearn.metrics import classification_report
4 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=.2, random_state= 33)
5 dect = DecisionTreeClassifier()
6 dect.fit(X_train, y_train)
7 print(classification_report(ytest, y_predict))
```



```
1 from sklearn.tree import DecisionTreeClassifier
2 from sklearn.model_selection import cross_val_score
3 result = cross_val_score(DecisionTreeClassifier(), X, y, cv=5)
4
5 result = [0.51376147, 0.32110092, 0.3853211 , 0.72222222, 0.84259259]
6 result.mean() = 0.5569996602106694
```

	precision	recall	f1-score	support
High	0.57	0.50	0.53	16
Low	0.80	0.77	0.78	26
Medium	0.57	0.60	0.59	20
Medium-High	0.47	0.78	0.58	9
Medium-Low	0.71	0.63	0.67	38
accuracy			0.65	109
macro avg	0.62	0.66	0.63	109
weighted avg	0.66	0.65	0.65	109



Discussion

1

Data Bising

2

3

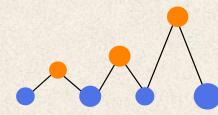
Feature Selection ,Varius Result and
Target Function Conversion

4

Overfitting Problem



Job Switching in Future and Reason



Experience Distribution Table

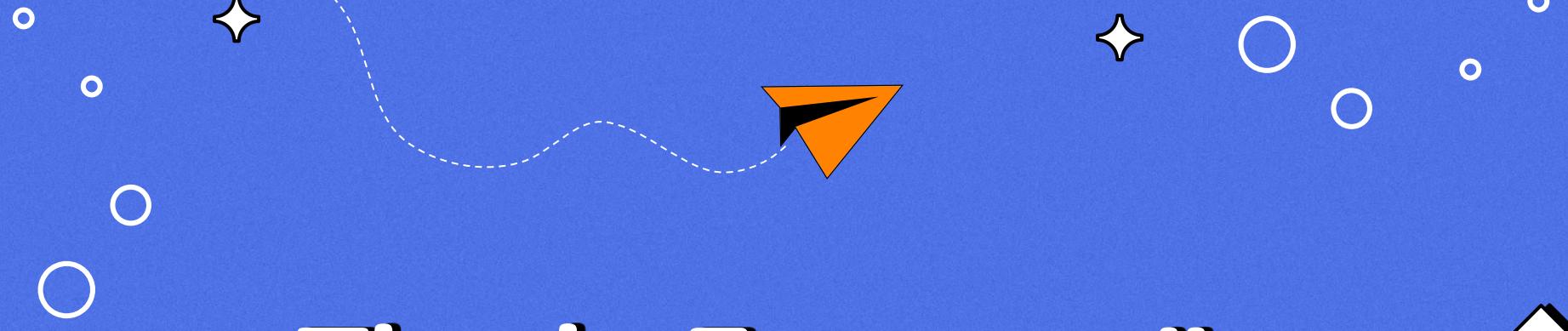
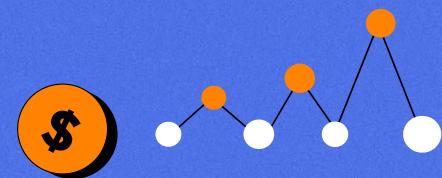
Required Experience	Count	Percentage(%)
Less than 2 years	119	21.92
More than 2 years	91	16.76
Fresh graduate	82	15.10
More than 3 years	65	11.99
More than 4 years	45	8.29
More than 5 years	39	9.21
Less than 1 years	36	6.63
More than 7 years	7	1.29
More than 6 years	6	1.10
More than 8 years	5	0.92

Required Experience	Count	Percentage(%)
More than 10 years	4	0.74



Thanks Everyone!!

Does Anyone have any question or Suggestion?



Desktop software

