INTERSHIP PROJECT

EMPLOYEE SALARY PREDICTION

BY - PAYAL BAJANTRI
PRN = 202401120064
BATCH = A3-44(DATA SCIENCE)

Overview

- Problem Statement
- System Development Approach
- Workflow
- Algorithm
- Deployment
- Results
- Conclusion
- Future Scope
- Reference

Problem Statement

In most organizations, salary prediction is still a manual process. This makes it time-consuming and sometimes unfair because it depends on subjective decisions. To solve this issue, I developed a system that can automatically predict the salary of an employee. The system uses two inputs, which are years of experience and job rate, and provides the output as the estimated salary. This makes the process faster, more reliable, and data-driven.

System Development Approach

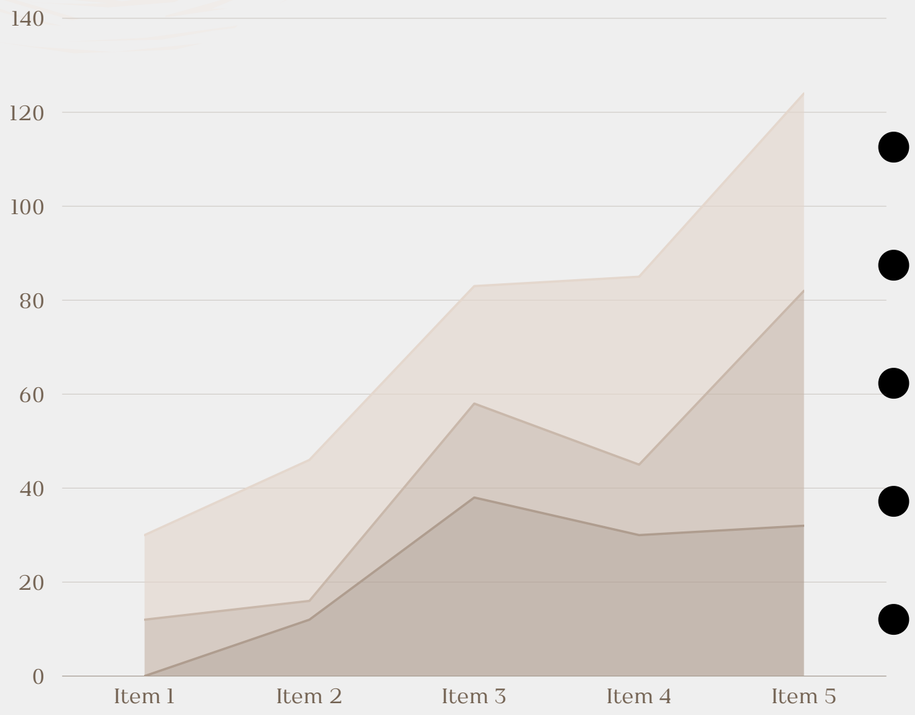
Hardware Requirement

- Normal laptop with 4 GB RAM
- Internet connection
- Libraries Used: NumPy, Pandas, scikit-learn, Joblib, Streamlit

Software Requirements:

- Python 3.8+
- Jupyter Notebook / VS Code
- Streamlit for web app

Workflow



- Collect dataset with Years of Experience, Job Rate, and Salary
- Clean and prepare the data
- Train a Linear Regression model
- Save the trained model using Joblib
- Build a Streamlit app with input fields
- Predict and display salary when button is pressed

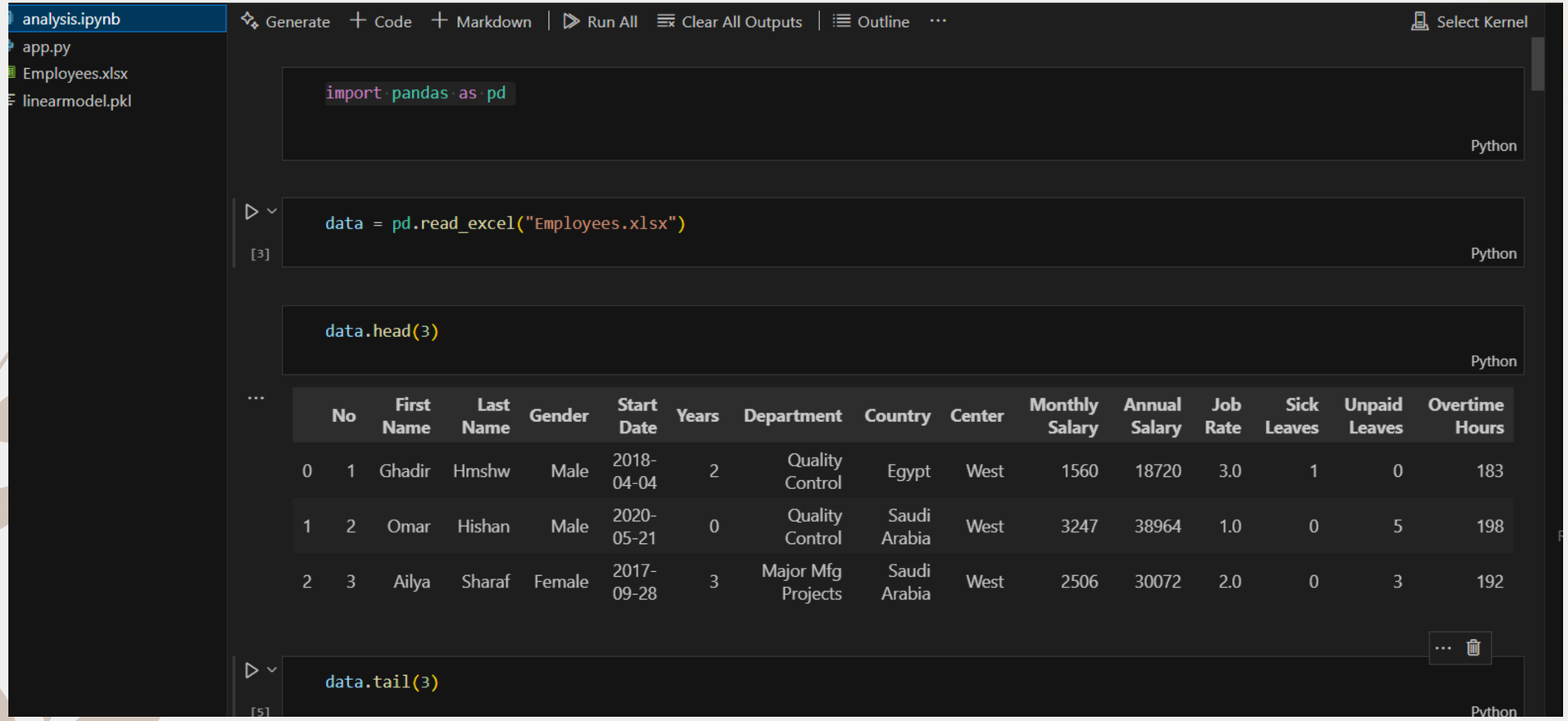
Algorithm (Linear Regression)

The algorithm used is Linear Regression. It is simple and effective for predicting numbers. Linear Regression works by finding a straight-line relationship between input and output. In this project, the inputs are years of experience and job

Deployment (Streamlit App)

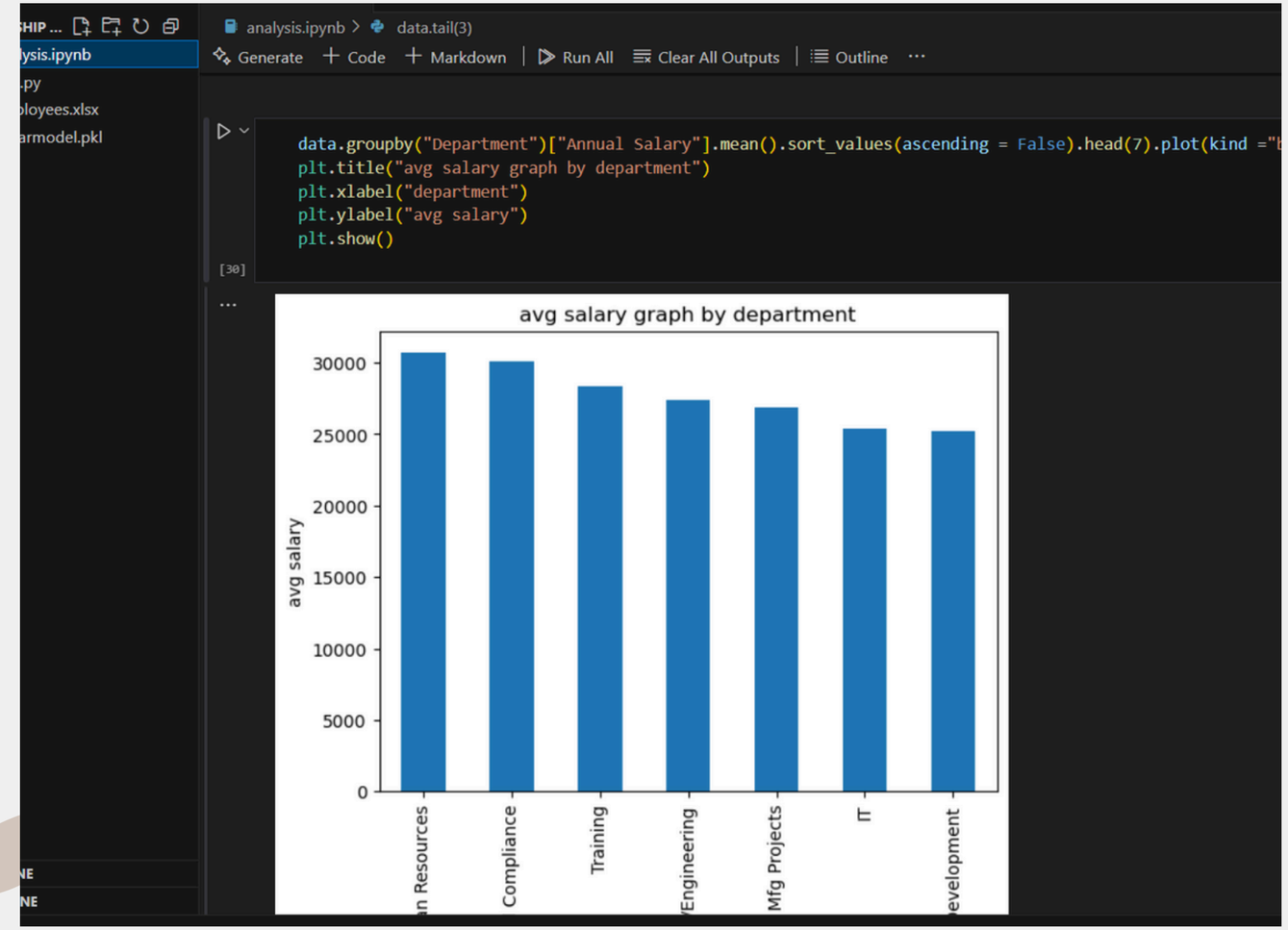
The app was created using Streamlit. Users can enter years of experience and job rate in the app. When the prediction button is pressed, the saved model runs in the background and the predicted salary is displayed instantly. The app is easy to use and can run both locally and online.

RESULT

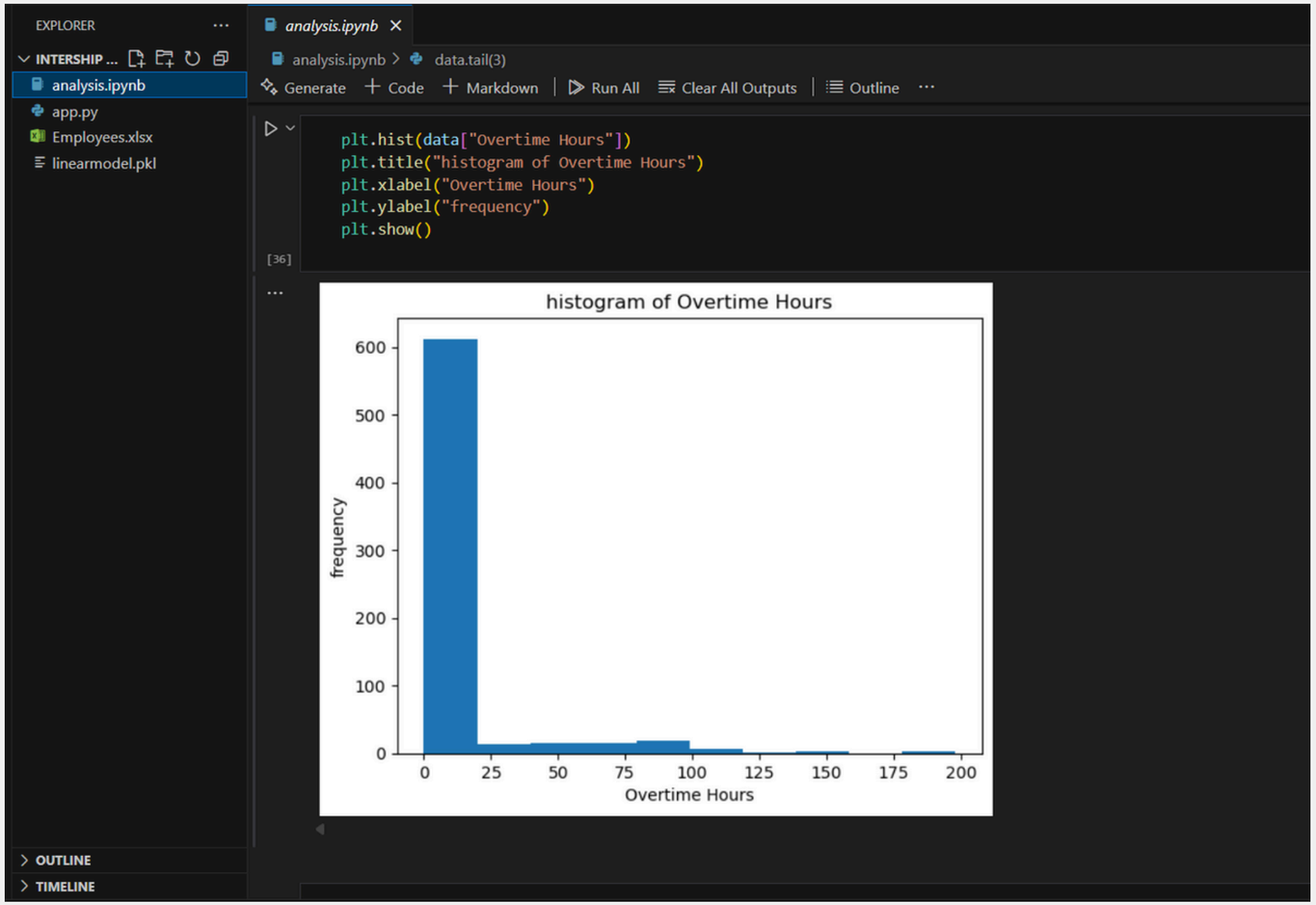


	No	First Name	Last Name	Gender	Start Date	Years	Department	Country	Center	Monthly Salary	Annual Salary	Job Rate	Sick Leaves	Unpaid Leaves	Overtime Hours
0	1	Ghadir	Hmshw	Male	2018-04-04	2	Quality Control	Egypt	West	1560	18720	3.0	1	0	183
1	2	Omar	Hishan	Male	2020-05-21	0	Quality Control	Saudi Arabia	West	3247	38964	1.0	0	5	198
2	3	Ailya	Sharaf	Female	2017-09-28	3	Major Mfg Projects	Saudi Arabia	West	2506	30072	2.0	0	3	192

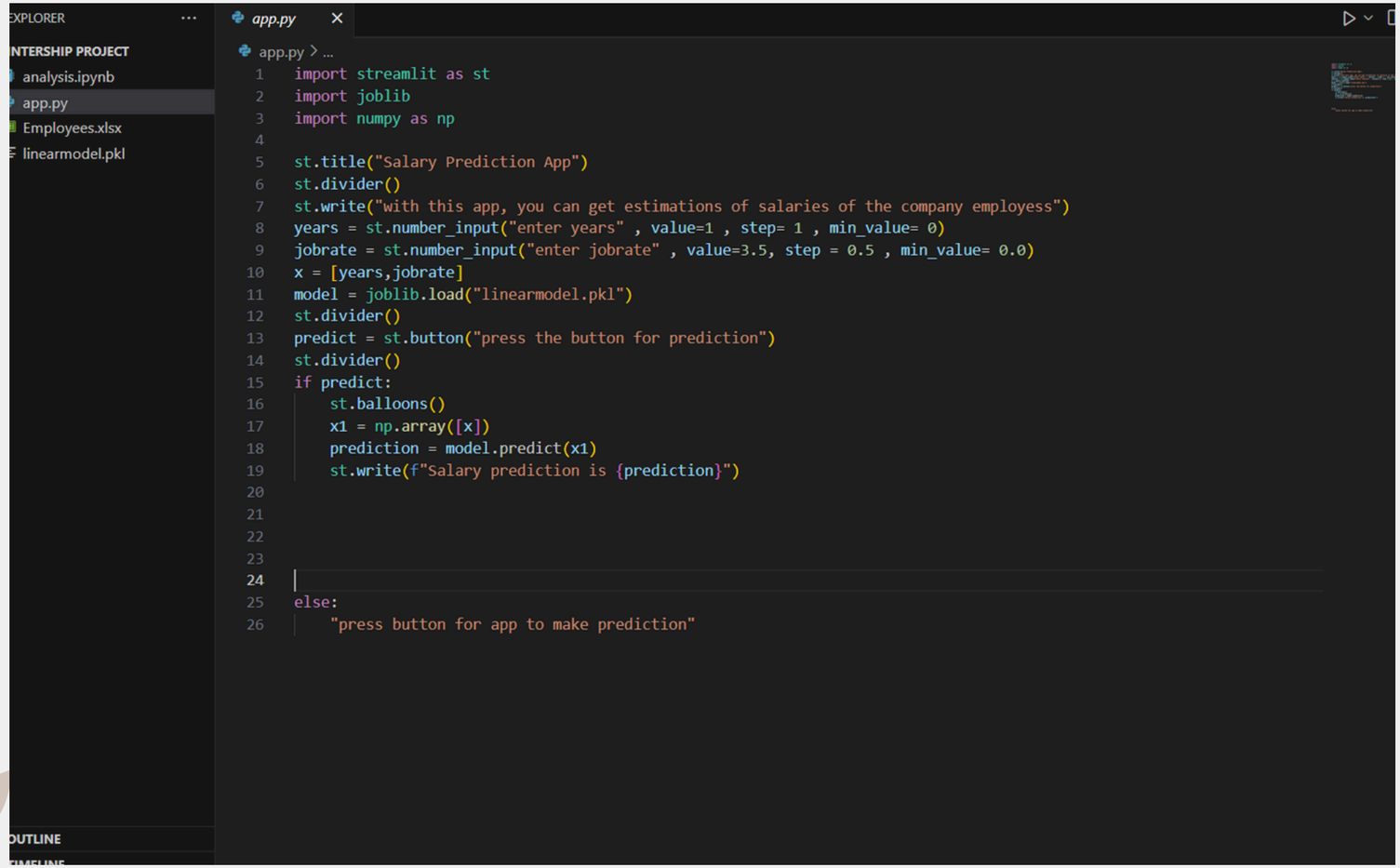
RESULT


analysis.ipynb > data.tail(3)
Generate
Code
Markdown
Run All
Clear All Outputs
Outline
ysis.ipynb
.py
loyees.xlsx
armodel.pkl
data.groupby("Department")["Annual Salary"].mean().sort_values(ascending = False).head(7).plot(kind ="b
plt.title("avg salary graph by department")
plt.xlabel("department")
plt.ylabel("avg salary")
plt.show()
[30]
avg salary graph by department
avg salary
30000
25000
20000
15000
10000
5000
0
Compliance
Training
Mfg Projects
IT

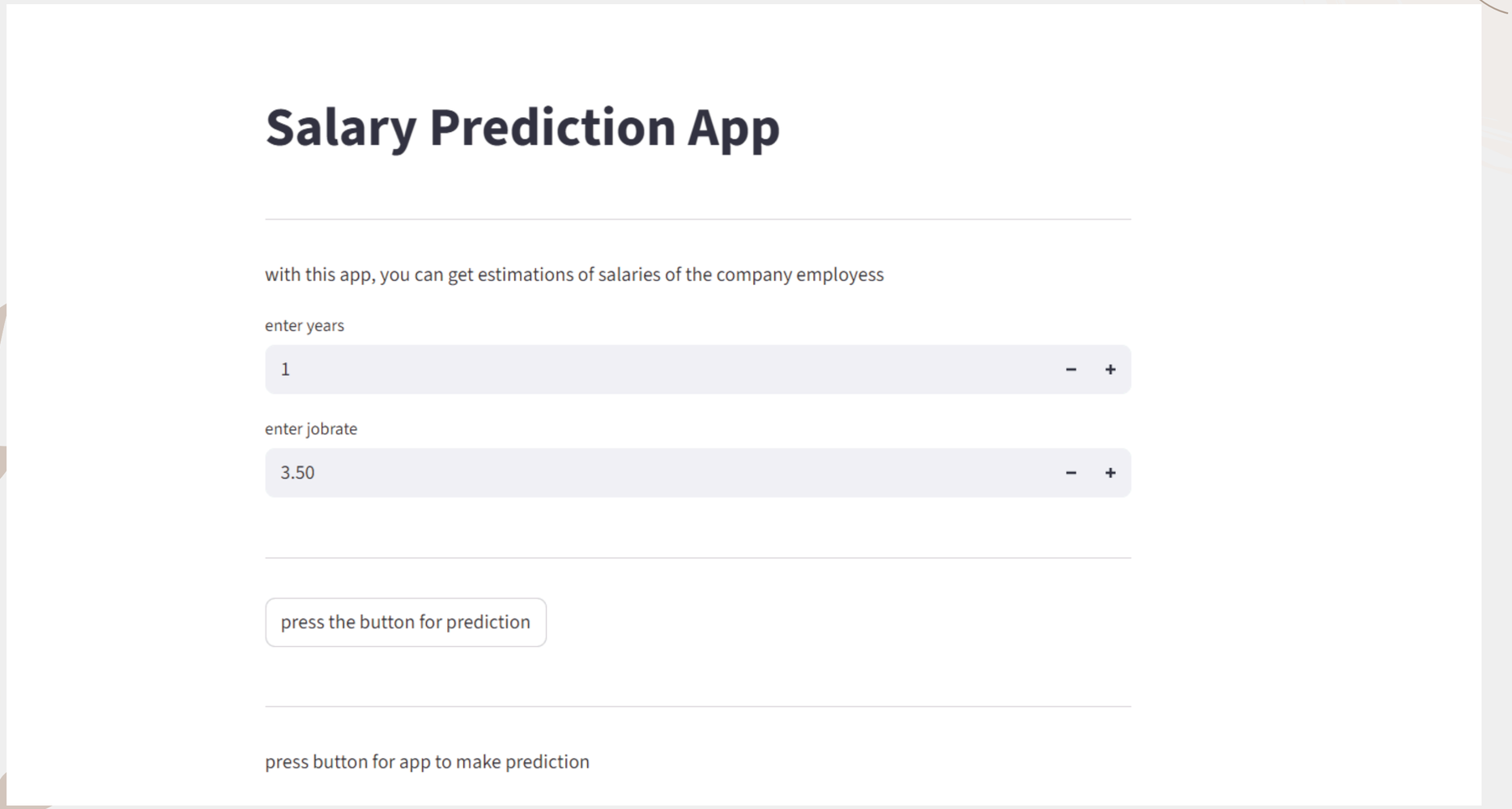
RESULT


EXPLORER
INTERSHIP ...
analysis.ipynb
app.py
Employees.xlsx
linearmodel.pkl
OUTLINE
TIMELINE
analysis.ipynb
analysis.ipynb > data.tail(3)
Generate
Code
Markdown
Run All
Clear All Outputs
Outline
plt.hist(data["Overtime Hours"])
plt.title("histogram of Overtime Hours")
plt.xlabel("Overtime Hours")
plt.ylabel("frequency")
plt.show()
[36]
histogram of Overtime Hours
frequency
Overtime Hours
0
100
200
300
400
500
600
0
25
50
75
100
125
150
175
200

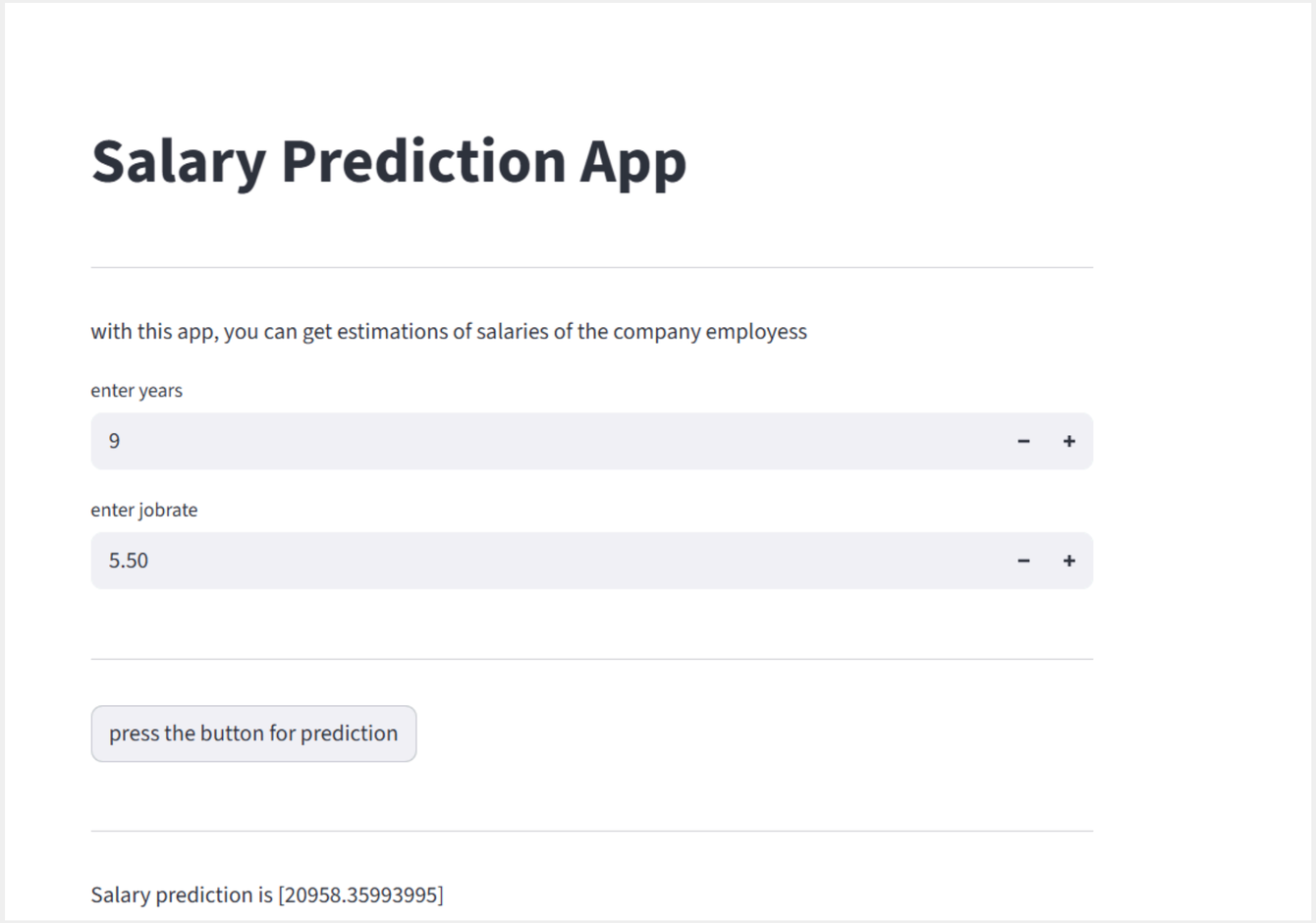
RESULT


EXPLORER
NTERSHIP PROJECT
analysis.ipynb
app.py
Employees.xlsx
linearmodel.pkl
OUTLINE
app.py
app.py > ...
1 import streamlit as st
2 import joblib
3 import numpy as np
4
5 st.title("Salary Prediction App")
6 st.divider()
7 st.write("with this app, you can get estimations of salaries of the company employess")
8 years = st.number_input("enter years" , value=1 , step= 1 , min_value= 0)
9 jobrate = st.number_input("enter jobrate" , value=3.5, step = 0.5 , min_value= 0.0)
10 x = [years,jobrate]
11 model = joblib.load("linearmodel.pkl")
12 st.divider()
13 predict = st.button("press the button for prediction")
14 st.divider()
15 if predict:
16 st.balloons()
17 x1 = np.array([x])
18 prediction = model.predict(x1)
19 st.write(f"Salary prediction is {prediction}")
20
21
22
23
24
25 else:
26 "press button for app to make prediction"

RESULT


Salary Prediction App
with this app, you can get estimations of salaries of the company employess
enter years
1
enter jobrate
3.50
press the button for prediction
press button for app to make prediction

RESULT


Salary Prediction App
with this app, you can get estimations of salaries of the company employess
enter years
9
enter jobrate
5.50
press the button for prediction
Salary prediction is [20958.35993995]

RESULT

The website gives correct predictions for given inputs. For example, if the input is 9 years of experience and a job rate of 5.5, the system shows the salary instantly. The predictions are realistic and useful for HR professionals as well as job seekers.

Conclusion

This project proves that machine learning can be applied in HR systems to make salary prediction simple and fast. By combining Python, Linear Regression, and Streamlit, I created a working system that is user-friendly and practical.

https://github.com/payalbajantri/internship.git

Future Scope

- Add more features such as education, job role, and location
- Use advanced algorithms to improve accuracy
- Deploy the app on cloud for global use
- Extend into a full HR analytics dashboard

Thank You So Much