CAPSTONE PROJECT

EMPLOYEE BURNOUT ANALYSIS

Presented By:
Saket Chaudhary
Veer Bahadur Singh Purvanchal University
B.Tech. (CSE)

edunet foundation

OUTLINE

- **Problem Statement** (Should not include solution)
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

- Employee burnout is a growing challenge in today's high-pressure work environments, characterized by physical, emotional, and mental exhaustion caused by prolonged stress. It adversely affects employee well-being, productivity, and job performance, leading to significant organizational setbacks. Identifying at-risk employees early is crucial but remains a complex task due to the interplay of various contributing factors such as workload, mental fatigue, and work-life balance. This project aims to analyze these factors and develop a predictive model to identify employees prone to burnout. By addressing this issue, organizations can take proactive measures to enhance workplace well-being and prevent burnout.

SYSTEM APPROACH

The "System Development Approach" outlines the strategy and methodology for developing and implementing the Employee Burnout Analysis project. It includes system requirements, libraries used, and the overall workflow for analysis and prediction.

System Requirements

- Hardware Requirements:
  - Processor: Intel Core i5 or above
  - RAM: 8GB or higher
  - Storage: At least 1GB free space for dataset and libraries
- Software Requirements:
  - Python 3.x environment
  - Jupyter Notebook or any compatible IDE

Libraries Required

The following Python libraries were utilized to develop the model:

1. *pandas*

- For data manipulation, cleaning, and preprocessing.

2. *matplotlib and seaborn*

- For visualizing trends, correlations, and key insights in the data.

3. *scikit-learn*

- For building and evaluating the predictive regression models.
- Submodules used:

  train_test_split for dataset splitting.

  StandardScaler for feature scaling.

  Regression algorithms like LinearRegression.

  Evaluation metrics like mean_squared_error and r2_score.

ALGORITHM & DEPLOYMENT

Algorithm Used:

- Regression Analysis: Predict burnout rates based on multiple employee metrics.
- Training the model on features like workload, mental fatigue, and work-life balance.

Deployment Process:

- Developed models in Jupyter Notebook.
- Evaluated using metrics:

  Root Mean Square Error (RMSE)

  Mean Absolute Error (MAE)

  R^2 Score for accuracy assessment.
- Visualizations for insights and interpretability.

RESULT

Dataset Loading and Overview:

Screenshot showing the dataset loaded in Jupyter Notebook, with key columns like "Mental Fatigue Score" and "Burnout Rate."

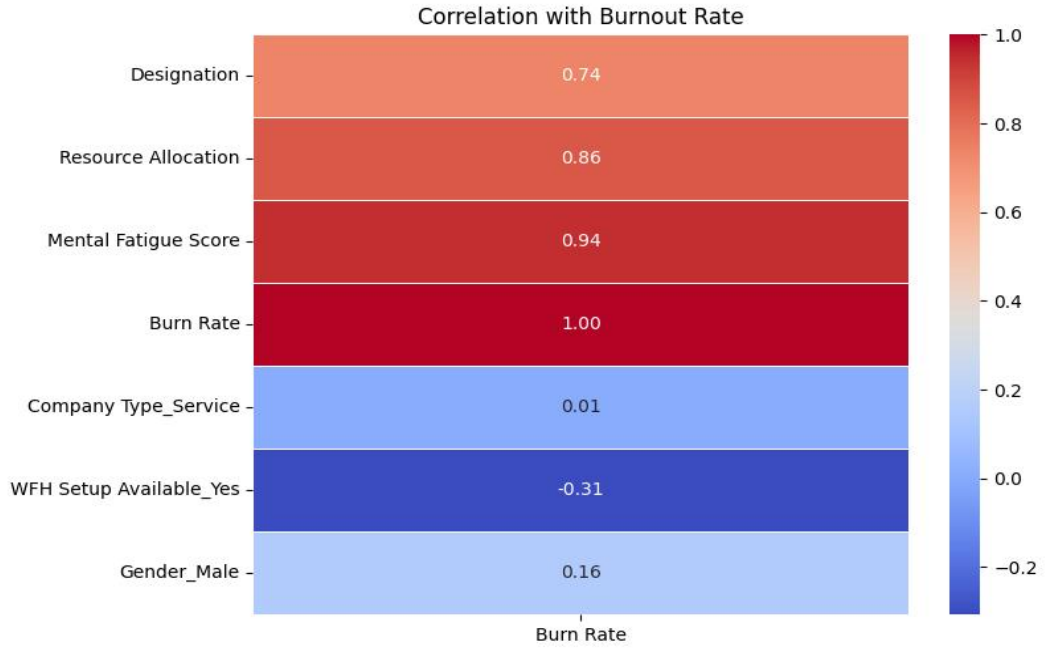
```
[ ] data = pd.read_excel ('/content/employee_burnout_analysis-AI 2.xlsx')
```

```
[ ] data.head()
```

	Employee ID	Date of Joining	Gender	Company Type	WFH Setup Available	Designation	Resource Allocation	Mental Fatigue Score	Burn Rate
0	fffe32003000360033003200	2008-09-30	Female	Service	No	2	3.0	3.8	0.16
1	fffe3700360033003500	2008-11-30	Male	Service	Yes	1	2.0	5.0	0.36
2	fffe31003300320037003900	2008-03-10	Female	Product	Yes	2	NaN	5.8	0.49
3	fffe32003400380032003900	2008-11-03	Male	Service	Yes	1	1.0	2.6	0.20
4	fffe31003900340031003600	2008-07-24	Female	Service	No	3	7.0	6.9	0.52

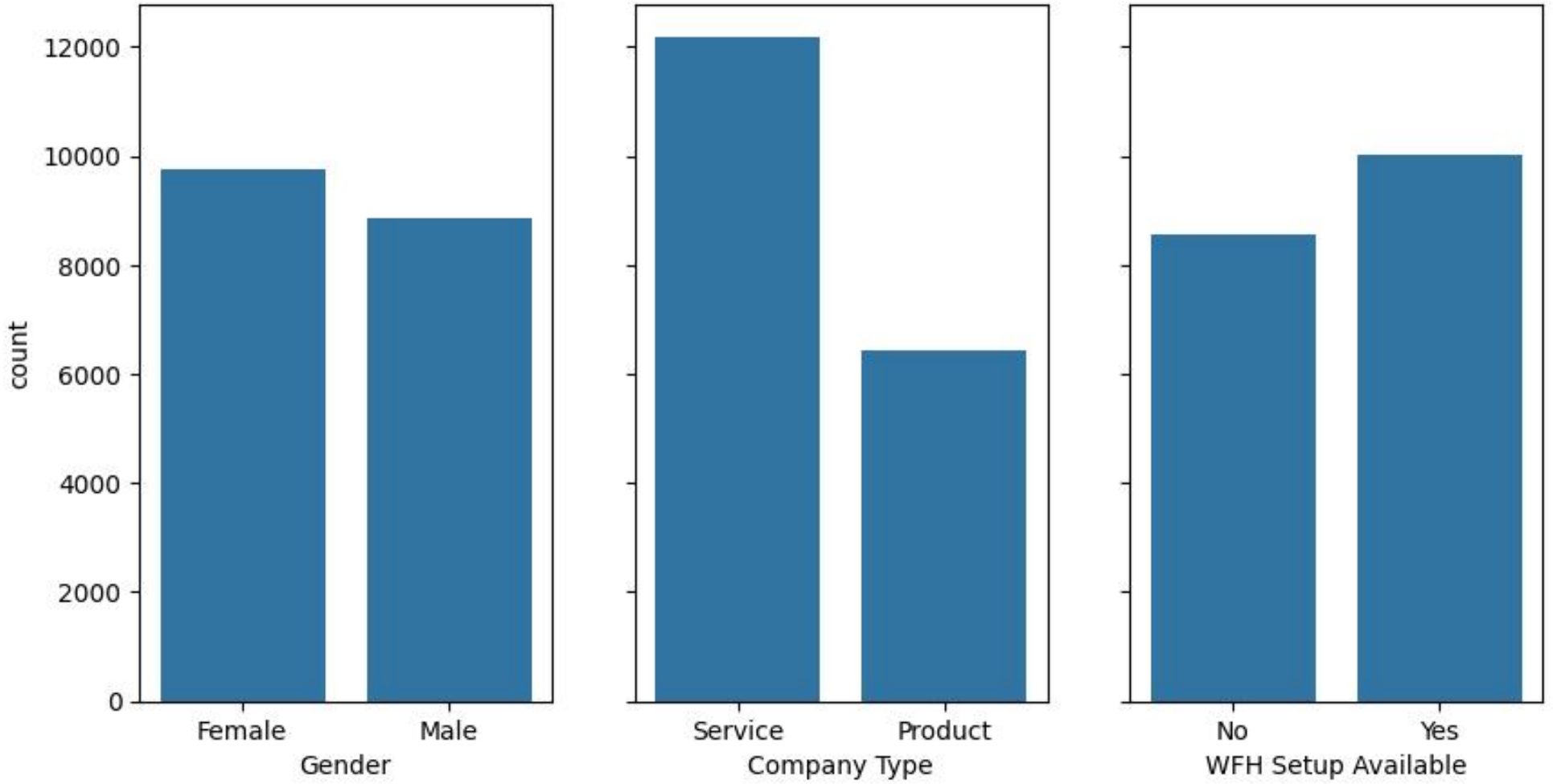
Exploratory Data Analysis:

Visualizations like the correlation heatmap highlighting the relationship between burnout rate and other features.



edunet foundation

Visual Representation of Categorical Data



Data Preprocessing:

Screenshot displaying missing value handling and feature encoding processes.

	Designation	Resource Allocation	Mental Fatigue Score	Company Type_Service	WFH Setup Available_Yes	Gender_Male
8977	0.723327	0.250185	-0.061773	0.724706	-1.082297	1.051505
14115	-0.159330	0.250185	-0.941481	0.724706	-1.082297	-0.951018
8797	0.723327	0.250185	0.973179	0.724706	-1.082297	-0.951018
1173	-1.041987	-1.214568	-0.579248	-1.379869	-1.082297	-0.951018
1941	-0.159330	0.738436	1.180169	-1.379869	0.923961	1.051505
...	...	...	...	...	...	...
13453	0.723327	1.226687	1.645897	-1.379869	0.923961	-0.951018
21179	0.723327	0.250185	-1.044976	0.724706	0.923961	1.051505
6327	0.723327	0.250185	0.093470	0.724706	-1.082297	1.051505
14933	-0.159330	0.250185	0.714441	0.724706	-1.082297	1.051505
288	-0.159330	0.250185	1.076674	-1.379869	-1.082297	-0.951018

13013 rows × 6 columns

Model Training and Evaluation:

Screenshot of the regression model training process, along with evaluation metrics like RMSE and R^2 Score.

```
Linear Regression Model Performance Metrics:

Mean Squared Error: 0.0031569779113610717
Root Mean Squared Error: 0.0561869905882231
Mean Absolute Error: 0.045950320326447 73
R-squared Score: 0.918822674247248
```

```
LinearRegression
LinearRegression()
```

edunet foundation

GitHub Link:

Employee Burnout Analysis Repository

You can access the full code, detailed analysis, and outputs through the provided repository link.

https://github.com/Saket22-CS/Employees-Burnout-Analysis.git

CONCLUSION

The Employee Burnout Prediction project successfully identified key factors like mental fatigue and workload as significant predictors of burnout. The regression model provided reliable forecasts, enabling organizations to proactively address employee well-being challenges.

- **Effectiveness**: The model demonstrated accuracy through metrics like RMSE and R^2, offering actionable insights to reduce burnout risks.
- **Challenges**: Issues with missing data and feature selection required iterative preprocessing and optimization.
- **Improvements**: Incorporating additional features, advanced algorithms, and real-time dashboards can enhance accuracy and scalability.

This project emphasizes the importance of data-driven solutions for fostering healthier and more productive workplaces.

FUTURE SCOPE

- **Integration with Real-Time Data:** Automate the model using live employee data from HR systems.
- **Feature Expansion:** Incorporate additional metrics like employee feedback, absenteeism, and productivity scores.
- **Improved Algorithms:** Experiment with advanced machine learning techniques (e.g., Random Forest, XGBoost).
- **Mobile/Web Integration:** Deploy a burnout prediction dashboard accessible to HR professionals.
- **Interactive Dashboards:** Develop user-friendly dashboards using tools like Streamlit or Power BI for real-time insights and reporting.
- **Cross-Industry Applicability:** Adapt the model for different industries, ensuring scalability and relevance across various organizational contexts.

REFERENCES

Dataset:

- Source: [https://www.kaggle.com/datasets/blurredmachine/are-your-employees-burning-out]

Python Libraries:

- pandas: https://pandas.pydata.org
- matplotlib: https://matplotlib.org
- seaborn: https://seaborn.pydata.org
- scikit-learn: https://scikit-learn.org

Additional Resources:

- Blog on Regression Analysis: https://towardsdatascience.com
- Machine Learning Documentation: https://docs.python.org

THANK YOU

edunet
foundation