Sri Sri University

Project - High Level Design

On

Employee Turnover Prediction Using Advanced Machine Learning

Team members:

Aditya Jyotiraditya Sahoo
Chinmaya Kumar Kar
Sthitapragyan Mahapatra

Guided by: Prof. (Dr). Pradipta Kumar Mishra
and
Prof. (Dr.) Rabinarayan Satpathy

Industry Mentor:
Sumit Shukla

Table of Contents

1. Introduction.....
    1.1. Scope of the document .....
    1.2. Intended Audience.....
    1.3. System Overview .....
2. System Design .....
    2.1. Process and Information Flow.....
3. Data Design .....
4. Non-Functional Requirements .....
5. References.....

Introduction

1.1. Scope of the document

This document contains High Level Design of the project Employee Turnover Prediction Using Advanced Machine Learning

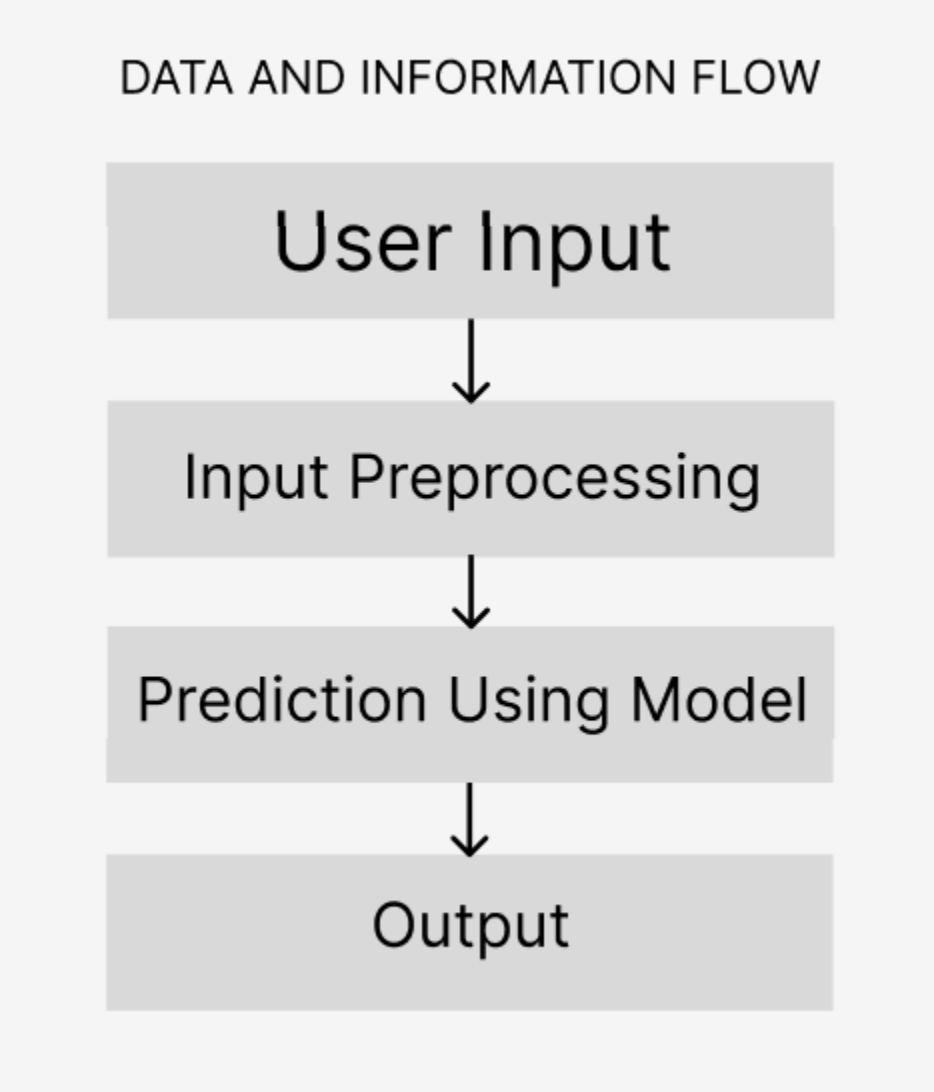
1.2. Intended Audience

The document is intended for team members of the project, faculty supervisor and industry mentor for the purpose of tracking and reviewing the progress.

1.3. System Overview

This project delves into the development of a powerful machine learning model for predicting employee turnover. By leveraging advanced algorithms, we aim to create an accurate and insightful model that can identify employees at risk of leaving. This model will analyze various factors that influence employee decisions, enabling organizations to proactively address potential causes of turnover. By providing early insights, the model can empower businesses to implement targeted retention strategies and foster a more engaged workforce, ultimately promoting long-term stability and growth.

2.1. Process Flow



Data Design

No. of Rows:- 14999

No. of Attributes:- 10

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14999 entries, 0 to 14998
Data columns (total 10 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   satisfaction_level     14999 non-null  float64
 1   last_evaluation        14999 non-null  float64
 2   number_project         14999 non-null  int64
 3   average_weekly_hours   14999 non-null  float64
 4   time_spend_company     14999 non-null  int64
 5   Work_accident          14999 non-null  int64
 6   left                   14999 non-null  int64
 7   promotion_last_5years  14999 non-null  int64
 8   department             14999 non-null  object
 9   salary                 14999 non-null  object
dtypes: float64(3), int64(5), object(2)
memory usage: 1.1+ MB
```

Non-Functional Requirements

Required Ram:- 6GB or higher
Required Storage:- 1GB or higher
Required Processor:- intel core i3 7th gen or higher
Required Libraries:-

- **numpy**
- **pandas**
- **seaborn**
- **matplotlib**
- **sklearn**

References

- NumPy –
- pandas - Python Data Analysis Library (pydata.org)
- scikit-learn: machine learning in Python — scikit-learn 1.4.1 documentation
- seaborn: statistical data visualization — seaborn 0.13.2 documentation (pydata.org)
- Matplotlib — Visualization with Python