📘 Employee Attrition Prediction at Salifort Motors

🔍 Objective

To analyze patterns in employee attrition and predict whether an employee is likely to leave the company using machine learning and deep learning techniques. The goal is to help HR proactively address factors influencing turnover.

📁 Dataset Description

The dataset includes 10,000+ rows with the following features:

Feature	Description
satisfaction_level	Employee satisfaction (0 to 1)
last_evaluation	Last performance score (0 to 1)
number_project	Number of projects completed
average_montly_hours	Monthly average working hours
time_spend_company	Tenure at the company (in years)
Work_accident	Whether the employee had a workplace accident
left	Target variable: 1 = left, 0 = stayed
promotion_last_5years	Was promoted in last 5 years (1 = yes, 0 = no)
Department	Department (e.g., Sales, IT, HR, etc.)
salary	Salary category (low, medium, high)

📊 Exploratory Data Analysis (EDA)

- **Attrition Rate:** ~24% of employees left the company.
- **Satisfaction Level:** Employees with low satisfaction are more likely to leave.
- **Tenure:** Attrition spikes at the 3–4 year mark.
- **Salary:** Most attrition is clustered among low-salary employees.
- **Department Trends:** Sales and support departments had higher turnover.
- **Work Accidents & Promotions:** No significant impact on attrition observed.

> Plots used: Correlation heatmap, box plots, bar charts, scatter matrix, pairplots.

🛠️ Feature Engineering

- **Tenure Categorization** (e.g., Junior, Mid, Senior)
- **Workload Ratio** = Monthly Hours ÷ Number of Projects
- **Performance Delta** = Evaluation Score − Satisfaction

All categorical features were encoded using Label or One-Hot Encoding, and numerical features were standardized.

🤖 Models Implemented

1. Traditional ML Models

- **Logistic Regression**: Baseline model with ~84% accuracy
- **Random Forest**: ~96% accuracy; high interpretability via feature importance
- **XGBoost**: ~96.7% accuracy; best performing classical model

2. Deep Learning

Feedforward Neural Network (FNN) using Keras:

- 2 hidden layers with ReLU activations
- Dropout for regularization
- Optimized using binary crossentropy and Adam optimizer

→ **Validation Accuracy: 96.7%+ after tuning with Keras Tuner**

✅ Outcomes & Recommendations

Insight	Recommendation
Low satisfaction, high work hours	Introduce wellness programs and workload balancing
High attrition in Year 3–4	Targeted interventions, career path planning
No promotion in 5 years = high attrition	Regular review for advancement opportunities
Sales & Support roles more vulnerable	Department-level retention strategies
Low salary = higher churn	Consider benchmarking and salary adjustments