**Employee Performance and Retention Analysis**

**1. Objective**

To analyze employee data and predict:

* **Performance Scores** using regression and deep learning.
* **Attrition Risk** using classification models.

**2. Key Findings & Insights**

**Performance Score Drivers:**

* **Salary**, **Years at Company**, and **Department** were the most influential predictors.
* Employees in **Engineering** consistently scored higher in performance evaluations.

**Attrition Risks:**

* Attrition was higher among employees with:
  + **Higher performance scores**
  + **Shorter tenure**
* Departments such as **Sales** and **Engineering** showed the **highest attrition rates**.

**3. Modeling Approach**

**A. Performance Prediction**

* **Linear Regression**:
  + Used as a baseline.
  + Evaluation metrics: **Mean Squared Error (MSE)** and **R² Score**.
* **Deep Learning with Keras**:
  + Architecture: Dense layers (64 → 32 → 1) with ReLU activation.
  + Loss function: **Mean Squared Error**.
  + Provided strong predictive accuracy and smooth learning curves.

**B. Attrition Classification**

* **Random Forest Classifier**:
  + Produced interpretable results with high precision and recall.
  + Evaluated using a **Confusion Matrix** and **Classification Report**.
* **Neural Network Classifier**:
  + Output layer: Sigmoid activation for binary classification.
  + Loss function: **Binary Crossentropy**.
  + Performance evaluated on **accuracy**, **confusion matrix**, and **classification report**.

**4. Visual Insights**

The project included multiple visualizations using Matplotlib, Seaborn, and Plotly:

* **Line Plot**: Average performance score by years at the company.
* **Bar Chart**: Attrition count by department.
* **Scatter Plot**: Salary vs. Performance Score.
* **Boxplot**: Distribution of performance scores across departments.
* **Model Validation**:
  + Actual vs. Predicted Performance Score.
  + Confusion Matrices for classification models.
  + Training history of loss across epochs.

**5. Recommendations**

**Performance Improvement**

* Implement **department-wise performance plans** (e.g., Sales coaching, cross-training in Support).
* Launch **mentoring and skill development programs** for early-tenure employees.
* Offer **performance-linked incentives** and recognition.

**Attrition Reduction**

* Focus retention strategies on:
  + **High-risk departments**
  + **Employees with low satisfaction indicators**
* Conduct **exit interviews**, **pulse surveys**, and implement **early warning alerts**.
* Support **career progression** and **internal mobility** to increase employee engagement.

**6. Tools & Libraries Used**

* **pandas**, **numpy** – Data handling and preprocessing
* **matplotlib**, **seaborn**, **plotly** – Visualization
* **scikit-learn** – Traditional machine learning
* **tensorflow.keras** – Deep learning models
* **scipy** – Hypothesis testing and statistical analysis

**Final Notes**

This project successfully combined statistical analysis, machine learning, and deep learning to uncover:

* Key performance drivers
* High-risk attrition segments
* Data-driven, department-specific recommendations

These insights are actionable for HR and leadership teams to **boost retention**, **enhance performance**, and **optimize workforce strategies**.