Project Title: AttriSight HR – Predictive Attrition & Engagement Dashboard with SHAP Insights

---

Project Summary:

AttriSight HR is a real-time, predictive HR dashboard that combines Google Sheets, Google Apps Script, and machine learning (trained via Google Colab) to visualize and explain **employee attrition risk**. Using SHAP (SHapley Additive exPlanations), the model also explains **why** an employee might leave, offering strategic insights for HR and leadership.

End-to-End Flow: From Raw Data to Predictive Dashboard

1. Data Collection (Google Sheets)

Raw HR data includes:

- Age, Tenure_Years, Engagement_Score, Performance_Rating, Manager_Rating, Promoted_Last_2Y, Department
- These are either historical data or live employee records maintained by HR teams.

2. Google Colab – ML Training & SHAP Logic

- **Model Used**: Logistic Regression / XGBoost

  Logistic Regression:

  Purpose: Predicts whether an employee will leave or stay based on factors like **Tenure**, **Engagement Score**, and **Manager Rating**.

  How It Helps: It identifies patterns in historical data to predict the likelihood of an employee's departure, helping HR take preventive actions.

  XGBoost:

  Purpose: A more advanced model that combines multiple decision trees to make accurate predictions. It handles complex relationships and larger datasets effectively.

  How It Helps: It looks at various factors like **performance**, **engagement**, and **promotion history** to give a precise risk prediction, especially when data is non-linear.

  SHAP - Making Predictions Transparent:

  What It Does: **SHAP** explains the contribution of each factor (e.g., **low engagement score** or **high tenure**) in the prediction.

  It helps HR understand **why** an employee might be at risk of leaving and provides actionable insights for **targeted interventions**.

Target: Attrition_Prediction (Yes/No)

Feature Engineering:

- Normalize tenure, bin engagement scores, one-hot encode departments, etc.

Explainability:

- Used **SHAP** to interpret the output of the model.
- SHAP values tell **which features drive attrition** for each employee.
- For instance:
  - Low Engagement Score (+ve SHAP → more likely to leave)
  - Promotion in last 2 years (–ve SHAP → likely to stay)

3. Pushing Output to Google Sheets

- After training, the model predicts:
  - Attrition_Risk_Percent
  - Attrition_Prediction (Yes/No)
- These columns are appended back to Google Sheets.
- SHAP summary plots and force plots can be exported as visuals or logs.

4. Google Apps Script + HTML Dashboard

- Pulls real-time data from Google Sheets.
- Offers 5 filters (Tenure, Engagement, Manager Rating, Risk Range, Department).
- Displays:
  - Risk by department
  - Risk level distribution
  - Performance & Engagement breakdown
  - **Radar Scatterplot** of employees' risk distribution
- Fully responsive and shareable via Web App.

Organizational Impact

Benefit	Description
Predictive Resignation	HR teams can identify high-risk employees *before* they exit.
Data-Driven Retention	SHAP insights help prioritize **why** someone may quit (e.g. poor engagement, low tenure).
Strategic Interventions	Targeted coaching, training, or compensation can be designed for high-risk employees.
Transparent HR AI	SHAP ensures the model's decisions are not a black box, improving **trust and fairness**.

💡 **What Makes This Unique?**

- **Google Sheets + Google Colab + SHAP**: Full-stack low-code HR analytics pipeline.
- **No external tools required**: Entire pipeline from model training to visualization stays in the Google ecosystem.
- **Modular and Scalable**: Can plug into HRMS like Workday or SAP via APIs