

Project 9 : Attrition Explainability with SHAP

Objective:

Enhance the HR Attrition prediction models by adding explainability using SHAP (SHapley Additive Explanations).

This notebook will:

- 1 . Load saved models from Project 7 (Deployment).
- 2 . Apply SHAP to explain feature contributions.
- 3 . Export explainability visuals for reports and dashboards.

Artifacts generated:

- SHAP summary plots
- SHAP bar plots
- SHAP force plots (employee-level explanations)
- Stored outputs in `/reports/09_explainability/`

Step 1 : Load Models & Data

We will use the Random Forest model trained in Project 7 and the same preprocessed data.

✓ Models & preprocessors loaded successfully.

Total trained columns: 44

i Detected raw dataset. Applying preprocessing...

✓ Data preprocessed successfully. Final shape: (1470, 44)

Step 2 : Initialize SHAP Explainer

We'll use SHAP's `TreeExplainer` since the model is a Random Forest.

Final aligned shape: (1470, 44) | Expected: 44

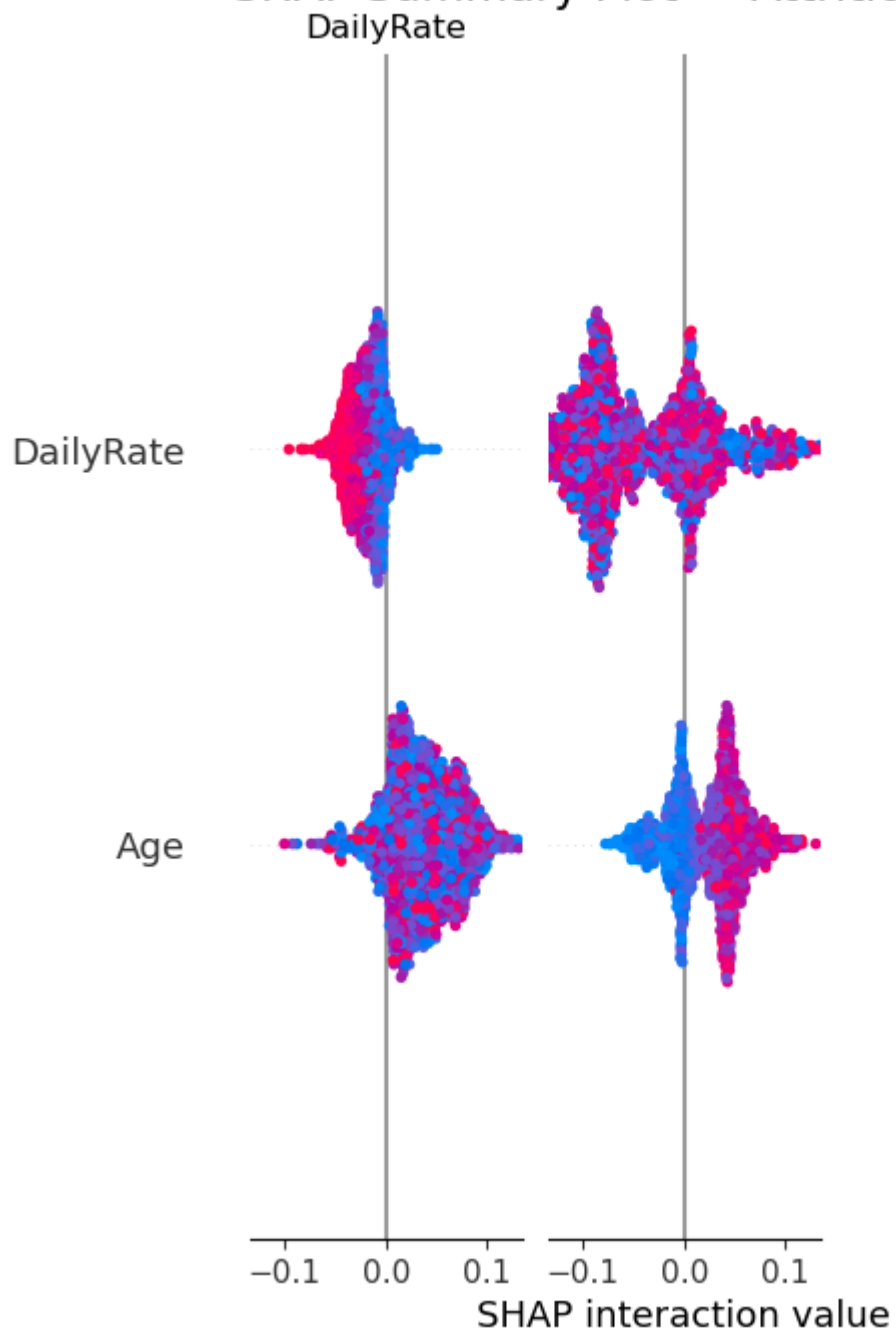
SHAP values computed. Shape: (1470, 44, 2)

Step 3 : Global Explainability — SHAP Summary Plot

This plot shows feature importance by their overall contribution to attrition predictions.

<Figure size 1200x800 with 0 Axes>

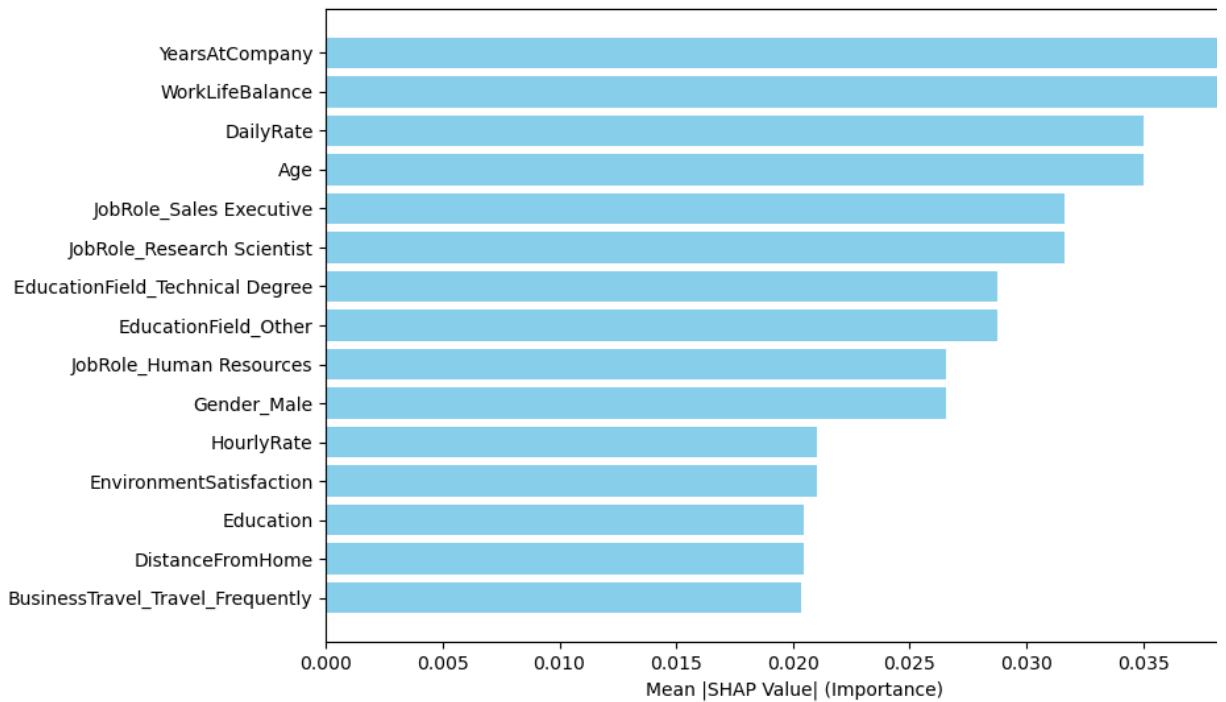
SHAP Summary Plot — Attrition Features



Step 4 : Global Explainability — SHAP Bar Plot

Rank features by average absolute SHAP value (importance).

SHAP Feature Importance (Top 15)



✓ Full importance table exported as CSV with all features.

Step 5 : Local Explainability — SHAP Force Plot

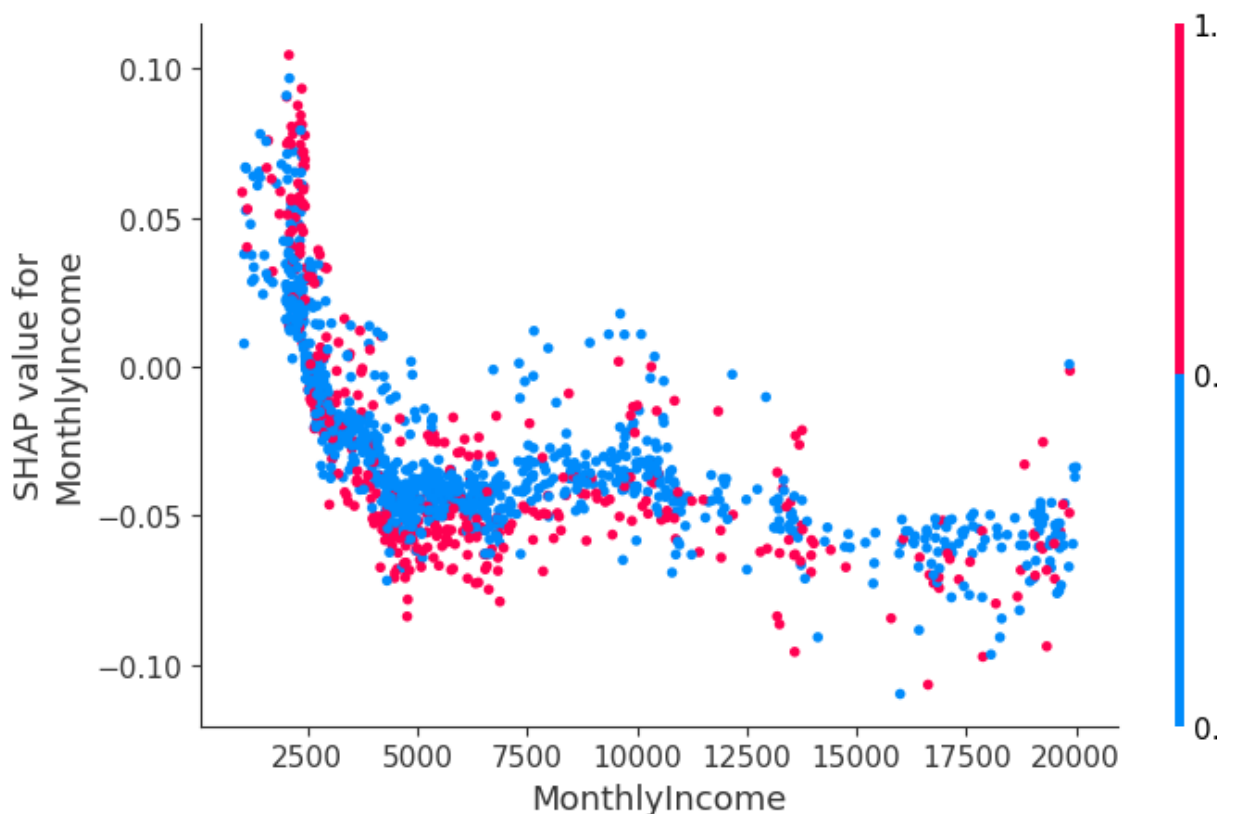
Explain **why** the model predicted attrition for one employee.

SHAP shape: (1470, 44) X_aligned shape: (1470, 44)

Aligned: perfect column match.

<Figure size 800x600 with 0 Axes>

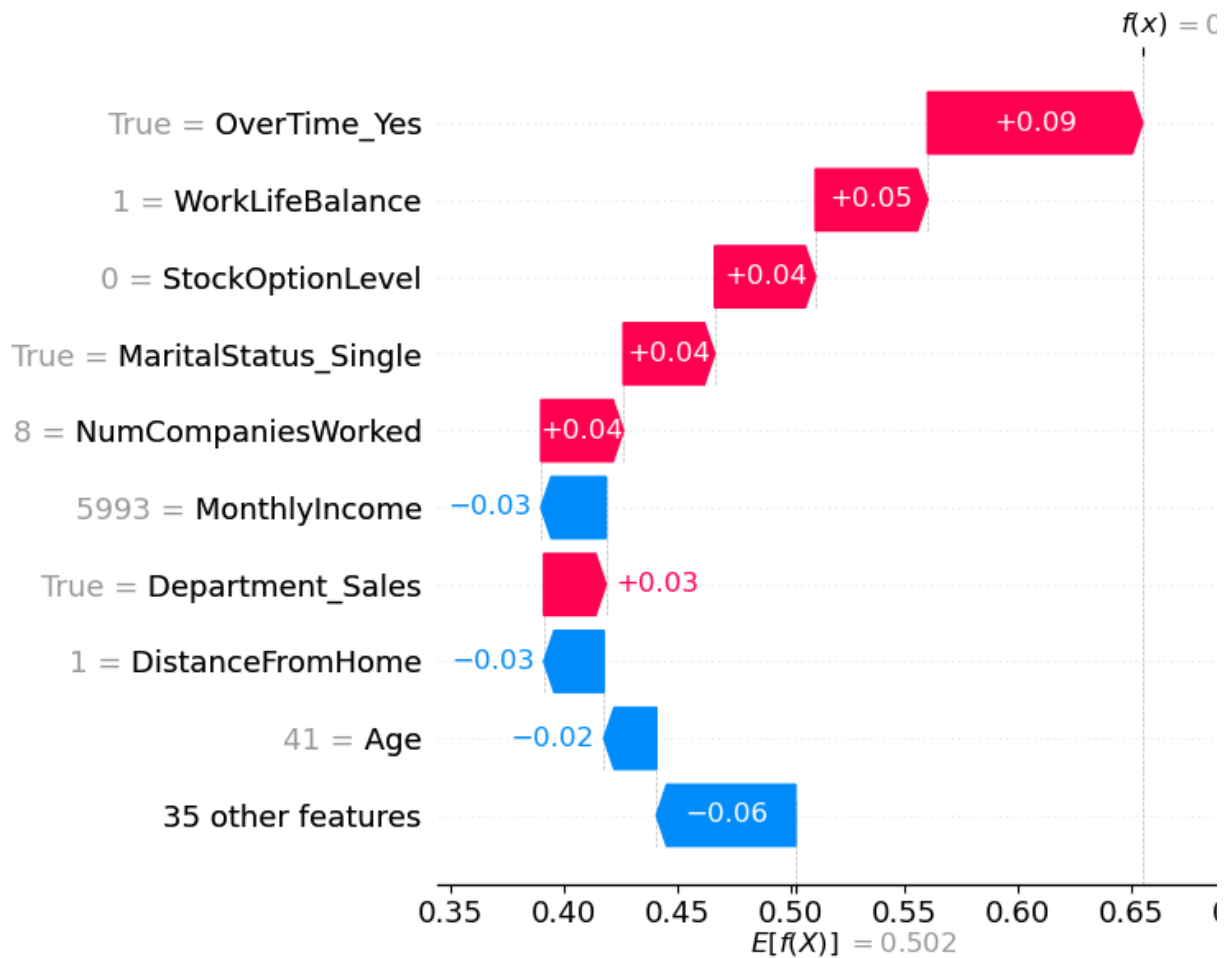
SHAP Dependence Plot — MonthlyIncome



✓ Dependence plot generated successfully.



SHAP Waterfall Plot — Employee 0



✓ Saved local explanation waterfall for Employee 0 -> images/shap_waterfall_employee0.png

✓ Conclusion

- SHAP confirmed the most important drivers of attrition.
- Features like **OverTime**, **MonthlyIncome**, **Age**, and **JobRole** consistently showed strong impact.
- These insights can now be integrated into the Streamlit dashboard for HR leaders.