Reducing Employee Attrition through Predictive Analytics

A Machine Learning Approach Using the IBM HR Dataset

PRESENTED BY: VISHNU LAL RAVEENDRAN NAIR

DATE: AUGUST 12, 2025

Table of contents

01. Executive Summary & Business Problem

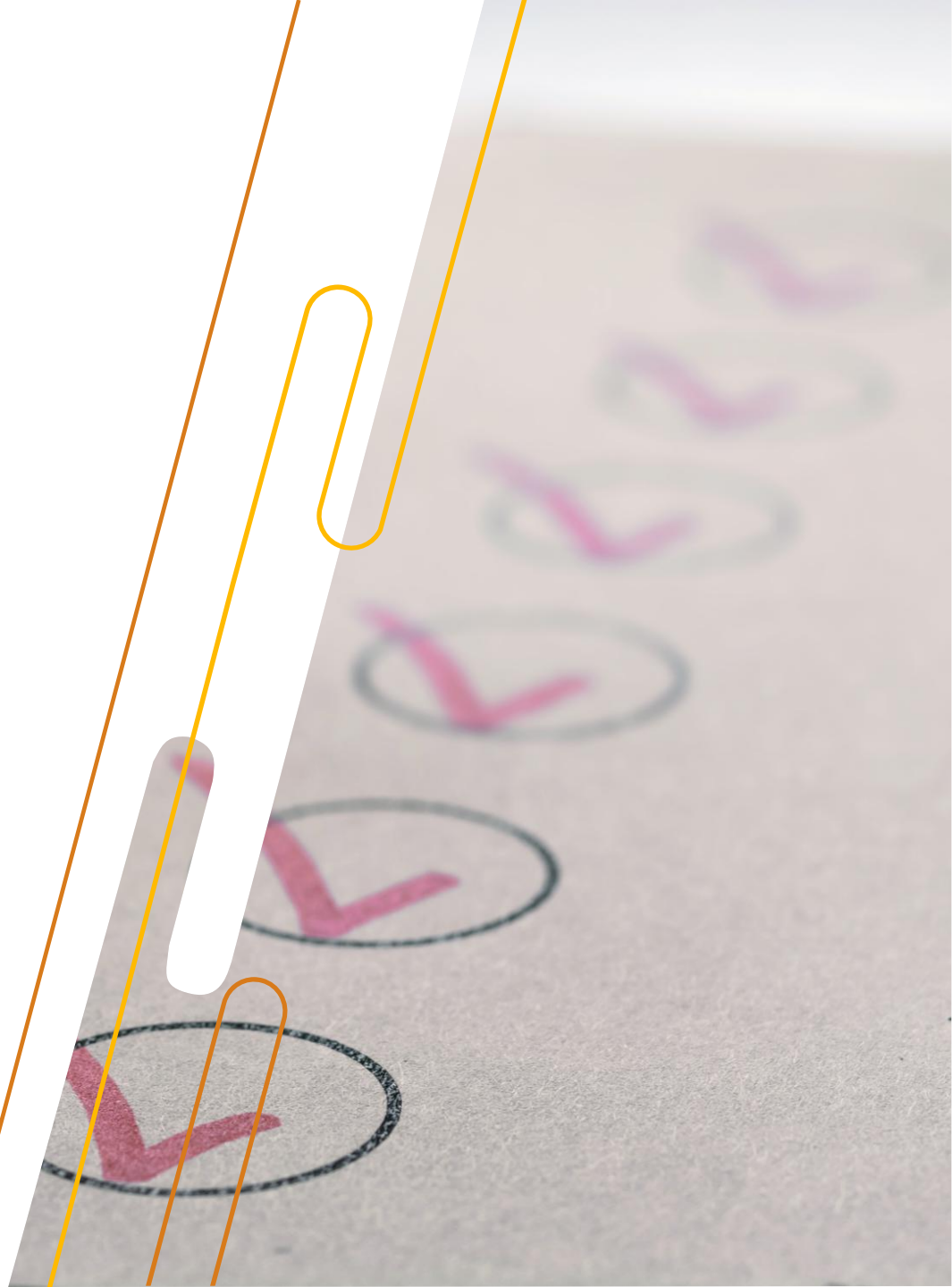
02. Project Objectives & Analytical Approach

03. Model Insights: What Drives Attrition

04. Employee Risk Profiling

05. HR Actions & Business Recommendations

06. Implementation Roadmap & Next Steps



Executive Summary

Purpose: Enable HR to proactively identify and retain at-risk employees

Approach:

- Combined internal HR data with external benchmarks
- Built 8 ML models, selected Logistic Regression
- Identified top attrition drivers using SHAP + odds ratio

Impact Potential: Up to **5–10% attrition reduction**

The Problem

Attrition is Costly and Hard to Predict

- 16% of employees left — hidden patterns across departments
- Replacement cost of **$4.8M annually – Annual Merit Increase budget.**
- Traditional dashboards show past leavers, not future risks
- Certain roles (e.g., Sales, frequent travelers) show disproportionate attrition

Data & Feature Highlights

- **Source:** IBM HR Analytics Attrition Dataset
- **Records:** 1,470 employees, 40 features
- **Target:** Voluntary Attrition (Yes/No)
- **No missing values**; complete, clean data

Category	Examples
Demographics	Age, MaritalStatus, Education, Gender
Job & Role Attributes	Department, JobLevel, JobRole, BusinessTravel, OverTime
Tenure & Movement	YearsAtCompany, YearsInCurrentRole, YearsWithCurrManager, YearsSinceLastPromotion
Compensation	MonthlyIncome, DailyRate, HourlyRate, MonthlyRate, StockOptionLevel
Engagement Drivers	JobSatisfaction, EnvironmentSatisfaction, RelationshipSatisfaction, WorkLifeBalance, JobInvolvement
Performance & Training	PerformanceRating, TrainingTimesLastYear

Engineered Feature	Definition	Purpose
CompaRatio	MonthlyIncome / Market Median for Role	Detect pay inequity and undercompensation
RoleTenureRatio	YearsInCurrentRole / YearsAtCompany	Identify career stagnation or late mobility
PromotionStagnation	YearsSinceLastPromotion / YearsAtCompany	Flag lack of progression over time
JobHoppingCategory	Binned version of NumCompaniesWorked adjusted for total experience	Reflect career stability patterns
PromotionFlag	1 if recently promoted (based on YearsSinceLastPromotion threshold)	Spot recently engaged talent

Modeling Summary

Modeling Goal
To build a classification model that:

- Accurately predicts voluntary attrition
- Is interpretable and HR-friendly
- Supports targeted retention actions based on real drivers

Evaluation Metrics

- **ROC AUC Score** (primary): Measures ability to separate classes across thresholds — handles imbalance well
- **F1 Score** (secondary): Harmonic mean of precision & recall — shows performance on the minority class (Attrition = Yes)

Model	Type	Rebalancing Used
Decision Tree (Basic)	Tree	None
Decision Tree (Tuned)	Tree	SMOTE + AUC Criterion
Random Forest	Ensemble (Tree)	SMOTE
XGBoost	Ensemble (Boosted)	SMOTE
Logistic Regression	**Regression**	**Class Weight (balanced)**
Logistic Regression (Stepwise)	Regression	Class Weights + AIC
Neural Network (Base)	Deep Learning	Class Weights
Neural Network (Tuned)	Deep Learning	Class Weights + Tuning

Model	ROC AUC	F1 Score	Interpretability	Business Readiness
Logistic Regression	**0.8362**	**0.5202**	✅ High	✅ Ready for HR use

LR: Top Drivers of Attrition

Feature	Mean_Abs_SHAP	Coefficient	Odds_Ratio	Interpretation
RoleTenureRatio	0.777863	-0.92412	0.396879	Longer time in same role relative to tenure reduces attrition
MonthlyIncome_log2	0.763516	-0.8986	0.407138	Higher income strongly reduces attrition risk
OverTime_Yes	0.761938	0.738623	2.093052	Overtime work more than doubles attrition risk
JobInvolvement_High	0.641516	-0.6604	0.516646	High involvement lowers attrition risk significantly
JobSatisfaction_Very High	0.622433	-0.72047	0.486524	Very satisfied employees are less likely to leave
YearsInCurrentRole_log2	0.609279	0.695007	2.003723	Absolute role duration increases likelihood of leaving
BusinessTravel_Travel_Frequently	0.565734	0.618872	1.856833	Frequent travel increases risk of exit
EnvironmentSatisfaction_High	0.556858	-0.57868	0.560639	High environment satisfaction lowers attrition
EnvironmentSatisfaction_Very High	0.54546	-0.66621	0.513654	Very high environment satisfaction reduces exits
WorkLifeBalance_High	0.528066	-0.5199	0.594579	Good work-life balance protects against attrition
EnvironmentSatisfaction_Medium	0.498205	-0.53907	0.58329	Medium satisfaction still lowers risk moderately
JobInvolvement_Medium	0.477009	-0.55567	0.573687	Moderate job involvement reduces risk slightly
JobSatisfaction_High	0.457738	-0.50811	0.601632	High satisfaction decreases likelihood of leaving
Department_Sales	0.445317	0.452674	1.572512	Employees in Sales face higher attrition risk
JobHoppingCategory_Moderate	0.44263	-0.45516	0.634344	Moderate job-hopping history reduces risk slightly

LR: Top Drivers of Attrition

Structural Drivers
These are job-related conditions that increase attrition risk:

Driver	Insight
Low Pay	**Below-market compensation** increases churn
Overtime	Employees working overtime are **2× more likely to leave**
Long time in same role	**Staying in the same role** for years raises risk
Role to Tenure Ratio	Role changes made too late don't reverse disengagement
Frequent Travel	High travel frequency linked to higher exits

Engagement & Satisfaction Drivers
These are experience-related factors protective against attrition

Driver	Insight
Job Involvement	Engaged employees are significantly more likely to stay
Job Satisfaction	Satisfaction scores strongly reduce attrition risk
Work-Life Balance	Better balance = lower turnover
Environment Satisfaction	Positive perceptions of workplace drive retention

Employee Risk Profiling

Feature	Mean_Abs_SHAP	Coefficient	Odds_Ratio	Interpretation
YearsInCurrentRole_log2	0.609279	0.695007	2.003723	For every **doubling of time** spent in the same role, the odds of attrition approximately double.
RoleTenureRatio	0.777863	-0.92412	0.396879	For every unit increase in RoleTenureRatio, the odds of attrition drop by 60%.

Scenario	Years in Current Role	Years at Company	RoleTenureRatio	Interpretation
Long time in role, moderate ratio	5	15	0.33	⚠️ **High risk**: Long-stayer in the same role — potential stagnation or disengagement
Long time in role, high ratio	8	8	1.00	🟡 **Medium risk**: Entire tenure spent in one role — loyal but may lack growth
Short time in role, short ratio	1	10	0.10	🟡 **Medium risk**: Recently moved after long tenure — may still be adjusting
Short time in role, high ratio	1	2	0.50	🟢 **Low risk**: Recent mover in early tenure — stable and engaged
Medium time in role, balanced ratio	4	7	0.57	🟢 **Low risk**: Healthy tenure and internal movement — no major red flags

- **RoleTenureRatio** = Years in role / Years at company
  - A **higher ratio** means an employee has spent **most of their tenure in the same role**
  - Strong **negative** impact on attrition (i.e., they are **less likely** to leave)
- **YearsInCurrentRole_log2** is **positively** associated with attrition (odds ratio > 2)
  - Long absolute time in role **may indicate stagnation**, leading to **higher attrition**

HR Recommendations

- **Rationale: YearsInCurrentRole_log2 – Odds Ratio: 2.003723 -** For every doubling of time spent in the same role, the odds of attrition approximately double. **RoleTenureRatio - Odds Ratio: 2.003723 - 0.396879 -** For every unit increase in RoleTenureRatio, the odds of attrition drop by 60%.

- **Recommendation 1: Address Career Stagnation with Proactive Mobility and Growth Paths (Odds ratio: 2.003723)**
  - **Identify stagnation risk:** Flag employees with ≥5 years in role and low RoleTenureRatio using the Power BI quadrant.
  - **Engage early:** Run stay interviews and career conversations for flagged employees.
  - **Promote mobility:** Offer stretch roles, cross-functional exposure, or internal transfers to re-engage and retain.

HR Recommendations

- **Rationale: MonthlyIncome_log2 – Odds Ratio: 0.407138** – For every doubling of monthly income, the odds of attrition decrease by approximately 59%,
- **Recommendation 2: Targeted Compensation Adjustments Based on Compa Ratio**
  - Instead of across-the-board salary increases, we propose a **targeted compensation strategy** based on **CompaRatio** — which compares an employee's salary to the market benchmark for their role and level.
  - Expected cost impact: 10.45M for correcting all, 4.1M for correcting below 75% Compa. Prioritize adjustments for high performers and those flagged as "High Risk" in our attrition model.
  - This ensures we address genuine pay gaps that influence attrition without inflating payroll costs unnecessarily.
  - Revisit Performance management system next year

HR Recommendations

- **Rationale: OverTime_Yes – Odds Ratio: 2.093052 –** Employees who work overtime have more than double the odds of attrition compared to those who do not.
- **Recommendation 3: Manage Overtime Risk Through Workload Controls and Compensation Equity**
  - Monitor and address excessive workloads
  - Incentivize fairly — especially when overtime isn't matched with competitive compensation.
  - Reduces burnout-related exits, ensures fair pay for effort, and sends a clear message that contribution is recognized and rewarded.

HR Recommendations

- **Rationale: BusinessTravel_Travel_Frequently – Odds Ratio: 1.856833** – Employees who travel frequently for work have nearly 1.9× the odds of attrition compared to those who travel rarely or occasionally
- **Recommendation 4: Address Business Travel-Linked Attrition with Tailored Actions for Sales vs. Non-Sales Roles**
  - In R&D and HR, excessive travel likely leads to burnout or disengagement. In Sales, travel is inherent — but the department also shows consistently high attrition, aligned with broader industry patterns
  - **For Sales:** Review and benchmark sales incentive plans, Adjust structures to drive performance and retention
  - **For Non-Sales Employees (e.g., R&D, HR):** Audit frequent travelers. Flag for manager review. Introduce travel balance guidelines or remote-first options

HR Recommendations

- **Rationale:** All engagement variables have negative coefficients and bring down attrition risk as they go up.
- **Recommendation 5: Leverage Engagement Survey Signals for Targeted Retention**
  - The model confirms that engagement factors — Job Involvement, Job Satisfaction, Work-Life Balance, and Environment Satisfaction — are strong protective drivers against attrition. Low scores in these areas increase attrition risk.
  - Flag employees with low scores as a priority group for retention attention
  - Align with model risk scores to focus retention strategies (stay interviews, team-level interventions)

Summary of Recommendations

#	Focus Area	Action Summary
1	**Pay Equity**	Prioritize compensation adjustments for employees with **CompaRatio < 75%/90%**
2	**Overtime Burnout**	Monitor frequent overtime; **adjust workloads** and review pay equity
3	**Career Stagnation**	Identify employees **≥5 years in role with low movement**; promote internal mobility
4	**Travel-Linked Risk**	Audit non-sales travel patterns; **review sales incentives and benchmark against market**
5	**Engagement Signals**	Use **survey data** (job satisfaction, involvement, WLB, environment) to flag risks

Implementation Roadmap & Next Steps

- **Phase 1: Immediate (0–1 month)**
  - Integrate model outputs into **Power BI dashboard**
  - Share insights with **HRBPs and department leads**
  - Conduct **stay interviews** for top 20 high-risk employees
- **Phase 2: Targeted HR Actions (1–3 months)**
  - Review compensation for employees with **CompaRatio < 75%/90%**
  - Launch workload reviews for **overtime-heavy teams**
  - Benchmark and update **Sales incentive programs**
  - Engage employees with **5+ years in same role**
- **Phase 3: Ongoing (Quarterly)**
  - Align **engagement survey** results with model risk scores
  - Track **attrition trends by department & risk group**
  - Add alerts for **new joiners, job hoppers, frequent travelers**

Thank you

Open to Questions!