# Promotion Prediction with AI — Results & Business Impact

Chris Stamou

September 2, 2025

## Objective

Build an explainable machine learning solution to predict which employees are likely to be promoted within 12 months, surface the key drivers, and replace a slow, subjective process with a repeatable, data-driven workflow. The data and the code are available on my GitHub repository https://github.com/ChrisStamou/AI-Promotion-Prediction.

## Synthetic Data & Visualization

We created a synthetic dataset of 5,000 employees representing a large professional services firm. The dataset includes demographic, performance, engagement, and organizational features, and was used to train and evaluate promotion prediction models.

To validate realism, we first visualized the data. Key plots include class balance, gender and department promotion rates, tenure and age distributions, and feature correlations. Figures are grouped below:

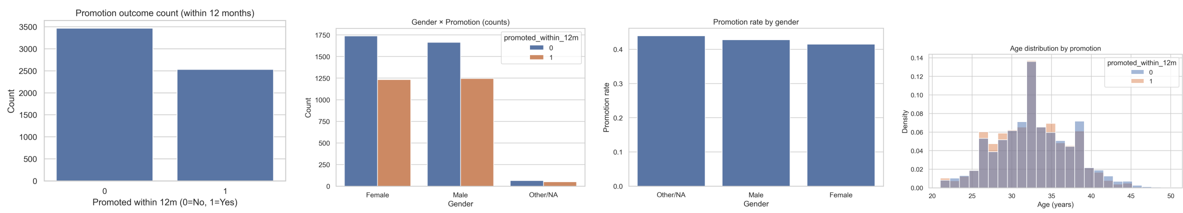


Figure 1: Exploratory plots (1/3). Promotion outcome balance, gender counts, gender rates, age distribution.

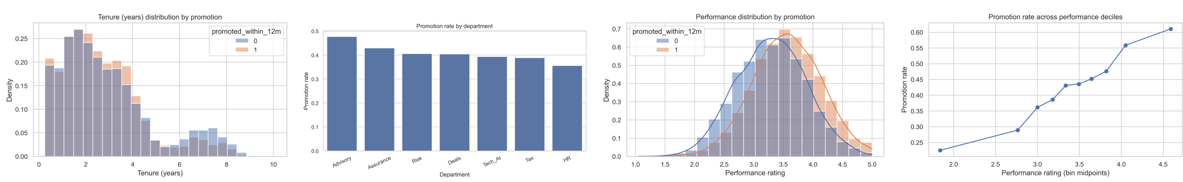


Figure 2: Exploratory plots (2/3). Tenure, department promotion rate, performance distribution and deciles.

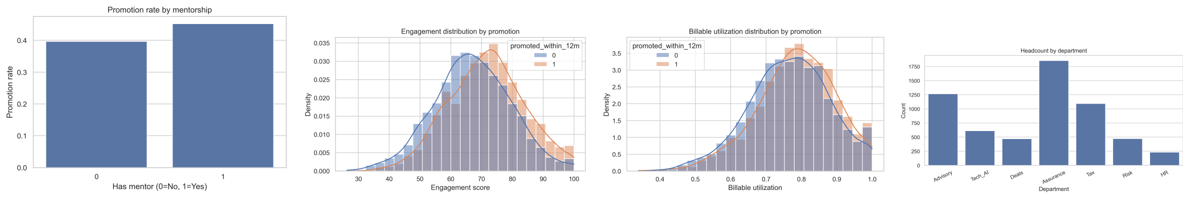


Figure 3: Exploratory plots (3/3). Mentorship, engagement, utilization distributions, headcount per department.

## Data & Method

- **Target:** `promoted_within_12m` (binary).
- **Features:** performance & manager ratings, tenure/eligibility, utilization, engagement, mentorship, department, location, education, etc.
- **Models:** Logistic Regression (baseline, interpretable), Random Forest (nonlinear), XGBoost (gradient boosting).
- **Evaluation:** Accuracy, Precision, Recall, F1, ROC AUC, PR AUC on a hold-out test set; ROC/PR curves and feature importance.

## Results

| Model | Accuracy | Precision | Recall | F1 | ROC AUC | PR AUC |
|---|---|---|---|---|---|---|
| Logistic Regression | 0.699 | 0.619 | **0.747** | **0.677** | **0.782** | **0.714** |
| Random Forest | **0.700** | **0.629** | 0.702 | 0.664 | 0.769 | 0.702 |
| XGBoost | 0.688 | 0.611 | 0.719 | 0.661 | 0.769 | 0.709 |

Table 1: Model performance on the test set.

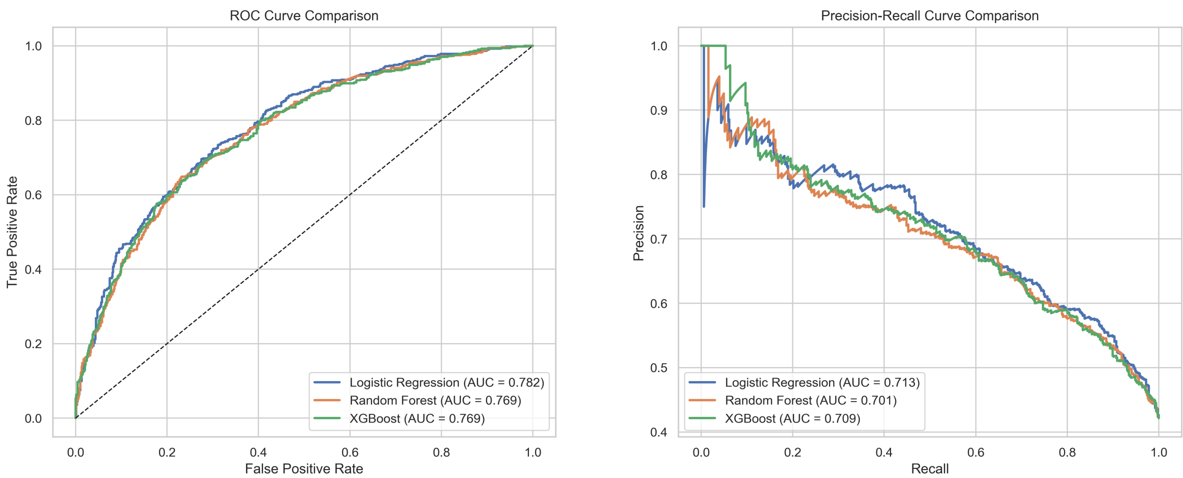


Figure 4: Combined ROC (left) and Precision–Recall (right) curves comparing models.

## Interpretation

All three models generalize well (ROC AUC ≈ 0.77–0.78). We recommend using **Logistic Regression** given its slightly higher AUC/recall and strong interpretability. Key drivers of promotion across models include: **eligibility/time since last promotion**, **performance**, **manager rating**, **engagement**, **billable utilization**, and **mentorship** (supportive effect; helps mitigate gender disparities).

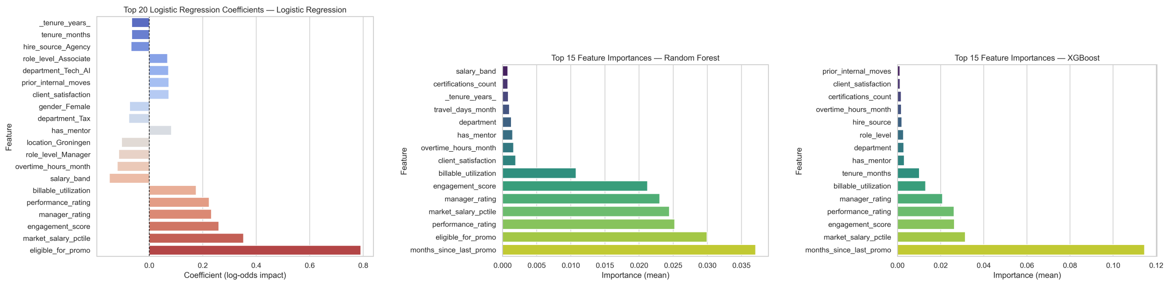


Figure 5: Feature importance visualizations: Logistic Regression coefficients, Random Forest importances, XGBoost importances.

## Example Employees

The tables below illustrate how HR can use the model-in this case we use XGBoost- to identify employees most and least likely to be promoted.

| ID | predicted_prob | actual | department | role | performance | engagement | mentor |
|---|---|---|---|---|---|---|---|
| 1117 | 0.933 | 1 | Advisory | Associate | 4.58 | 70.2 | 0 |
| 1626 | 0.925 | 1 | Advisory | Associate | 4.08 | 84.6 | 0 |
| 3112 | 0.911 | 1 | Assurance | Associate | 4.21 | 98.9 | 1 |
| 3656 | 0.910 | 1 | Advisory | Associate | 4.71 | 81.6 | 1 |
| 1335 | 0.900 | 1 | Advisory | Sr_Assoc | 4.43 | 78.0 | 0 |
| 480 | 0.890 | 1 | Assurance | Associate | 4.66 | 71.0 | 1 |
| 468 | 0.883 | 1 | Tax | Associate | 4.59 | 100.0 | 0 |
| 2152 | 0.881 | 1 | Tax | Associate | 4.27 | 85.2 | 1 |
| 1979 | 0.881 | 1 | HR | Associate | 4.68 | 83.6 | 0 |
| 2173 | 0.880 | 1 | Advisory | Sr_Assoc | 4.70 | 74.6 | 0 |

Table 2: Top 10 likely promotions (model prediction).

| ID | predicted_prob | actual | department | role | performance | engagement | mentor |
|---|---|---|---|---|---|---|---|
| 3117 | 0.015 | 0 | Tax | Associate | 3.67 | 50.7 | 0 |
| 845 | 0.019 | 0 | Advisory | Associate | 2.69 | 72.6 | 0 |
| 4658 | 0.020 | 0 | Tax | Associate | 2.07 | 61.5 | 1 |
| 3403 | 0.022 | 0 | Assurance | Associate | 3.27 | 59.9 | 0 |
| 4536 | 0.023 | 0 | Deals | Sr_Assoc | 3.59 | 79.6 | 0 |
| 5194 | 0.024 | 0 | Advisory | Associate | 2.26 | 61.1 | 0 |
| 5511 | 0.025 | 0 | Assurance | Associate | 2.86 | 50.1 | 1 |
| 2486 | 0.027 | 0 | Tech_AI | Associate | 3.52 | 64.2 | 0 |
| 4105 | 0.029 | 0 | Assurance | Associate | 2.78 | 67.9 | 1 |
| 4258 | 0.029 | 1 | Assurance | Associate | 3.48 | 66.6 | 1 |

Table 3: Bottom 10 unlikely promotions (model prediction).

## Business Impact

**Time-Saving Impact of AI in Promotion Analysis**

Promotion analysis is an essential task for HR because it directly impacts workforce planning, attrition, and financial forecasting. In a traditional setup without AI, HR would rely on manual reporting and consultations with managers across multiple departments. While this approach can save some time compared to reviewing each employee individually, it remains slow, resource-intensive, and heavily dependent on subjective judgments. Moreover, when the process is rushed due to the scale of the workforce, the results risk being less credible and not truly data-driven.

By contrast, with an AI-assisted approach, the same task can be executed in hours rather than weeks, while ensuring consistency, transparency, and evidence-based insights. Instead of spending over a month coordinating reviews and manually identifying patterns, HR can use the model to automatically detect promotion drivers, forecast likely promotions, and flag potential risks. This represents a realistic efficiency gain of thousands of hours saved per year, while also improving the quality and reliability of the insights delivered to leadership.

### Quantifying the Time Savings (Illustrative)

**Scope:** 5,000 employees in a large professional services practice.

**Manual, manager-led process (blended effort):**

- Data pulls, cleaning, slicing: $\sim 35-45$ hours
- Department reviews & calibrations (prep, meetings, notes): $\sim 120-140$ hours
- Pattern-finding & manual analysis across drivers: $\sim 40-50$ hours
- Deck/report writing, revisions with Finance/Leadership: $\sim 20-25$ hours

**Total (manual):** $\approx 215-260$ hours $\Rightarrow$ **1.3–1.6** FTE months (at 160 h/month).

**AI-assisted workflow:**

- Run modeled pipeline on latest HR CSV (training/scoring): $\sim 1-2$ hours
- Validation/QA with HRBPs & spot-checks by department: $\sim 16-24$ hours
- Report refresh (plots/tables/summary): $\sim 6-8$ hours

**Total (AI-assisted):** $\approx 23-34$ hours $\Rightarrow$ **0.15–0.2** FTE months.

**Net time saved per cycle:** $\approx 190-230$ hours (**85–90**% reduction).
Run semi-annually, this yields $\approx 380-460$ hours saved per year (**2.4–2.9** FTE months), while improving consistency and explainability.

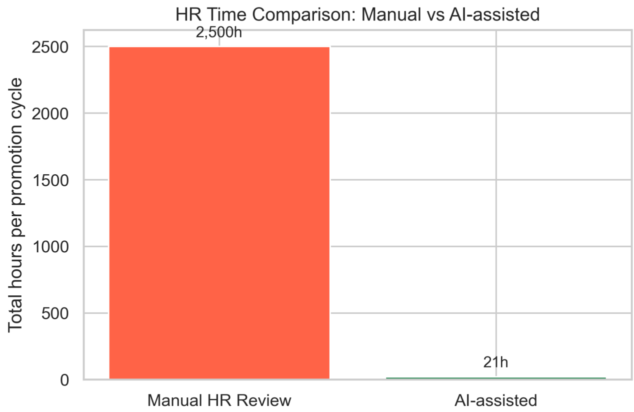


Figure 6: Manual vs AI-assisted HR time (illustrative). Manual $\approx 215-260$ hours vs AI-assisted $\approx 23-34$ hours per cycle.

*Note:* If HR chose a purely per-employee manual review (e.g., 0.50 hours per employee $\times$ 5,000 = 2,500 hours), the effort would be much higher. In practice, teams shorten this via manager consultations; our blended estimate reflects that more realistic—but still time-consuming—approach, which AI materially compresses.

**Planning value:** The model enables promotion volume forecasting by department, budget impact assessment (salary/bonus), and early fairness checks. The workflow is repeatable: re-run every 6–12 months with an updated HR CSV export.

## Take-home Message

Promotion prediction is a critical element of workforce planning. It allows HR to anticipate talent needs, align leadership pipelines, and manage financial implications of promotions such as salary increases and bonuses. Without AI, this process is manual, subjective, and time-intensive, often relying on manager opinions and lengthy reviews.

By leveraging AI, the same analysis can be performed in hours rather than weeks. The models not only predict who is most likely to be promoted, but also reveal the underlying drivers such as eligibility, performance, engagement, and mentorship. This transforms promotion planning into a data-driven, transparent, and repeatable process.

**In short:** Promotion prediction with AI empowers HR to make faster, fairer, and more efficient decisions. It reduces manual effort, enhances credibility of insights, and ensures that leadership decisions are supported by evidence rather than intuition.