LOGISTIC REGRESSION ON EMPLOYEE PROMOTIONS

Predictive Analytics for Talent Management

Tanya Gampert, PHR, CAPM
WGU – MSDA – Decision Process Engineering

Artificial Intelligence and HR: The Regulatory Environment

THE REGULATORY ENVIRONMENT
LAWS

6 states have AI laws on the books regarding talent management as of 2026, and the list is growing.

Source: https://www.brightmine.com/us/resources/hr-compliance/ai-laws-by-state-and-locality/

State	Bill Number	Requirements
California	OAL Matter Number 2025-0515-01 and 11 CCR § 7120	Impact and bias assessment for automated decisions
Colorado	2024 Bill Text CO S.B. 205	Bias audits
Illinois	2023 Bill Text IL H.B. 3773, amending the Illinois Human Rights Act (IHRA)	Prove non-discrimination
Maryland	2020 Bill Text MD H.B. 1202	Transparency
New York (NYC)	NYC Administrative Code § 20-871	Bias audits
Texas	2025 Bill Text TX H.B. 149	Prove non-discrimination

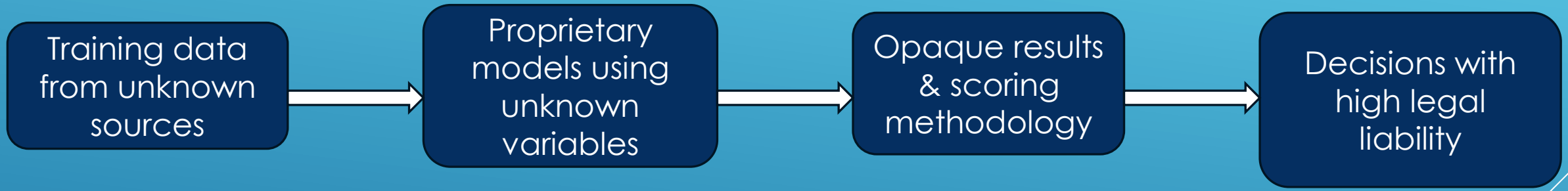
THE REGULATORY ENVIRONMENT
LAWSUITS

Organizations need to ensure that decision algorithms are transparent and unbiased.

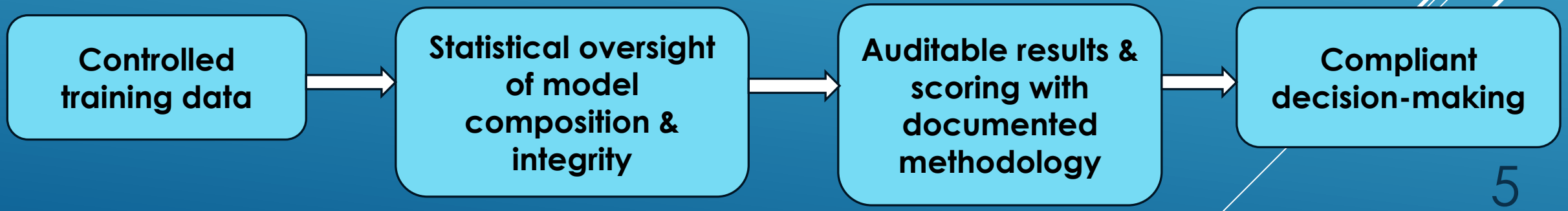
Source: https://www.nytimes.com/2026/01/21/business/ai-hiring-tools-lawsuit-eightfold-fcra.html

THE REGULATORY ENVIRONMENT
AI BLACK BOX VS. INTERNAL MODELS

AI Black Box



Internal Models



The Analysis

THE ANALYSIS
THE RESEARCH QUESTION

Can logistic regression accurately predict employee promotion outcomes for use in talent decision optimization?

Null hypothesis (H_0): Logistic regression cannot accurately predict promotions.

Alternate hypothesis (H_a): Logistic regression can predict promotions with accuracy above 80%.

THE ANALYSIS
LOGISTIC REGRESSION

Why logistic regression?

- Well-established statistical technique for binary classification
- Easy to interpret and maintain
- Reproducible and transparent insights over black box algorithms

Promotion probability = Employee promotion readiness relative to other employees

What are the chances of **being promoted** versus **not being promoted**?
What predicts the chances of **being promoted**?
To what extent do the chances of **being promoted** differ among **employees**?

THE ANALYSIS
MODEL CREATION

Number of EEs (new hires excluded): **9,271**
Promotion Rate: **17%**

Variables kept

Field
role_level
education_level
years_in_role
performance_rating
awards
kpis_achieved_pct
peer_review_score
training_courses_completed
certification_count
mentorship_participation
projects_delivered
romoted (target)

The Analysis Process:

- Cleaned 10,000 EE records by excluding Directors & new hires
- Tested model assumptions
- Handled class imbalance through threshold optimization
- Used coefficient analysis and model metrics to interpret results

Variables removed

Field	Drop Reason
department	Statistically insignificant/irrelevant to merit
years_in_company	Proxy of years_in_role
kpis_count	Proxy of role_level
salary	Proxy of role_level
performance_intervention	n = 30/irrelevant to merit

THE ANALYSIS
LOGISTIC REGRESSION RESULTS

$$\ln\left(\frac{Predicted\ Probability\ of\ Promotion}{1 - Predicted\ Probability\ of\ Promotion}\right) = -1.9381$$

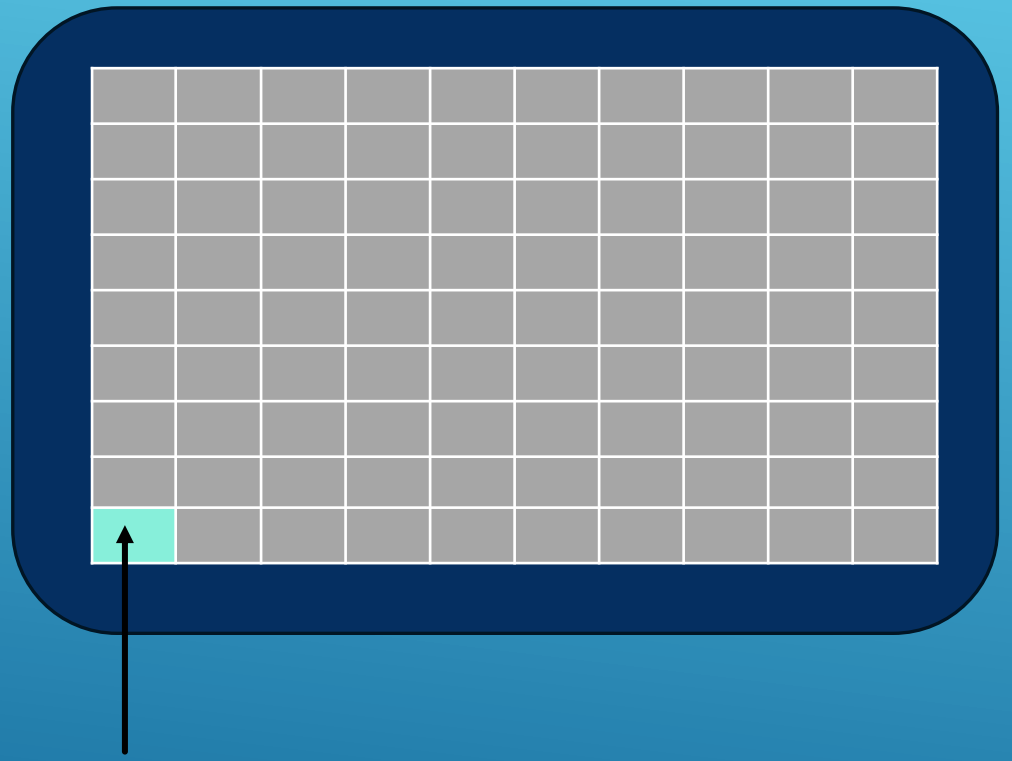
$$-0.0055 * \textbf{\textit{Years in Role}}$$
$$-0.0159 * \textbf{\textit{Performance rating}}$$
$$+0.0975 * \textbf{\textit{Awards}}$$
$$-0.0055 * \textbf{\textit{KPIs Achieved Percent}}$$
$$+01575 * \textbf{\textit{Peer Review Score}}$$
$$+0.206 * \textbf{\textit{Training Courses}}$$
$$+0.0021 * Certification\ Count$$
$$-0.0049 * \textbf{\textit{Projects Delivered}}$$
$$+0.0040 * Role\ Level$$
$$+0.0660 * Education\ Level$$

The negative intercept:
When all variables are considered, the chances of promotion go down.

Strongest indicators of promotion:
- Awards
- Peer review score
- Training courses

Negative coefficients – factors working against promotion:
Years in role, Performance rating, KPIs achieved, Projects delivered

THE ANALYSIS
MODEL RESULTS IMPLICATIONS



???

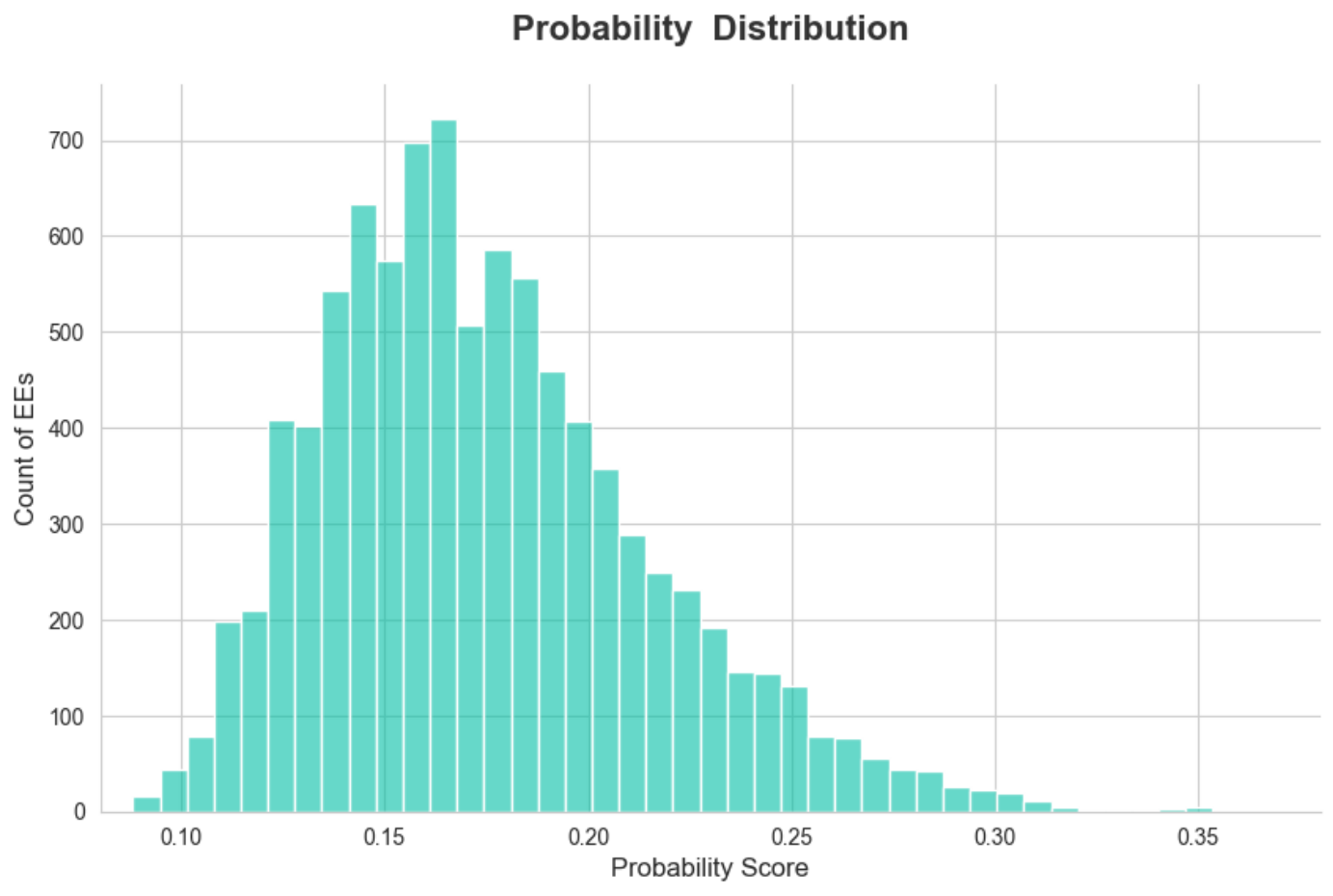
Unknown variables not included in the merit dataset account for **98.8%** of the variance in promotions.

Whatever is driving promotions is mostly not captured in the provided merit variables.

The variables in the merit dataset only explain **1.2%** of the variance in promotions.

Pseudo $R^2 = 0.012$

THE ANALYSIS
MODEL OUTPUT



The highest probability of getting promoted is 0.35, representing the top of the scoring range.

The middle of the range is 0.17.

The lowest probability score is 0.088, representing the bottom of the range.

THE ANALYSIS
SUMMARY OF FINDINGS

Can logistic regression accurately predict employee promotion outcomes for use in talent decision optimization?

This study failed to reject the Null Hypothesis, but the model still produced usable results for the purpose of talent classification.

Observations:

- The model has low accuracy due to few actual promotions and shadow variables, but can be improved with additional data
- There is a disconnect between the merit-based process vs. the actual promotion process
- Employees are penalized for merit achievements in current promotion decisions

Limitations:

- Synthetic data does not capture politics or relationships
- Not enough explainability from the variables in this dataset

Operationalizing the Model

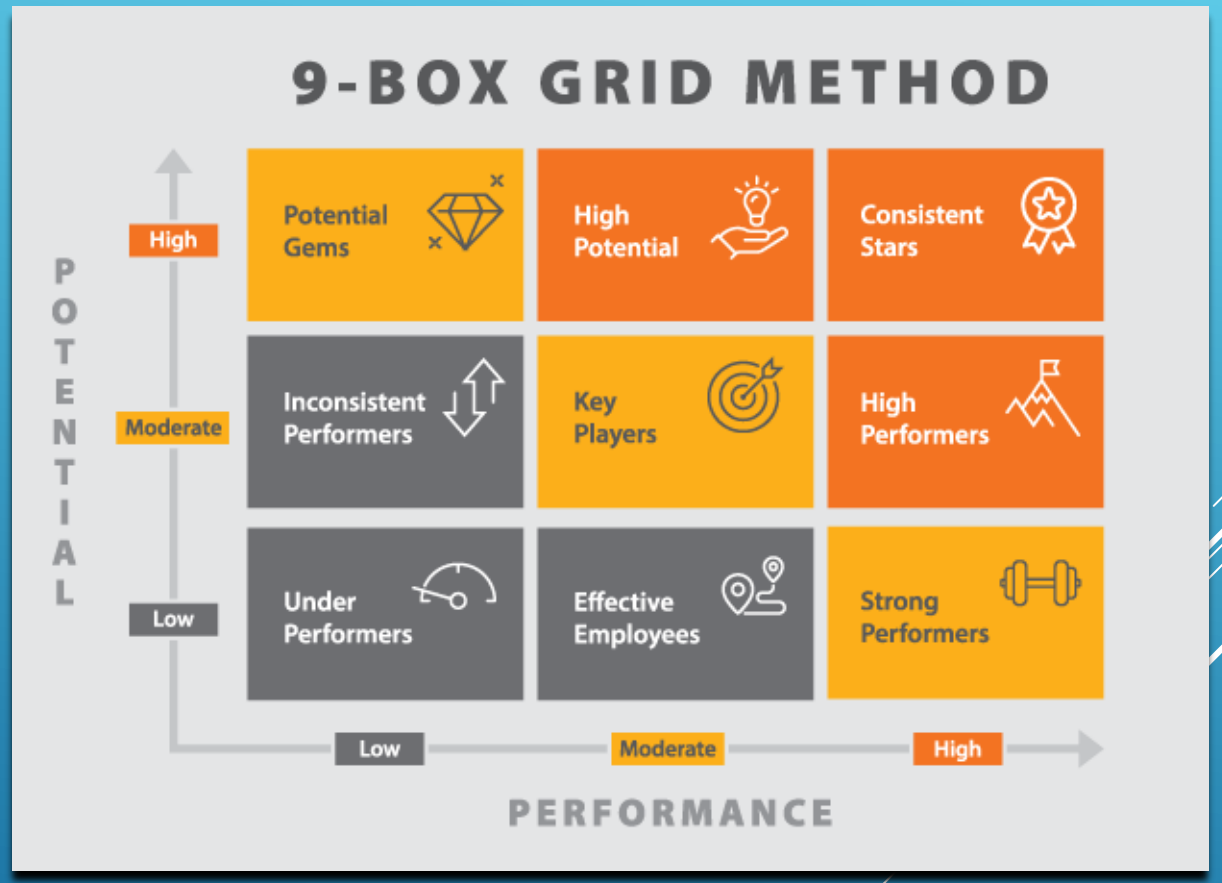
OPERATIONALIZING THE MODEL
9 BOX CLASSIFICATION METHODOLOGY

POTENTIAL
3 tiers of probabilities:

1. High = Highest 25%
2. Moderate = Middle 50%
3. Low = Lowest 25%

PERFORMANCE
3 tiers of subjective performance ratings:

1. High = 4+
2. Moderate = 3
3. Low = 1 or 2

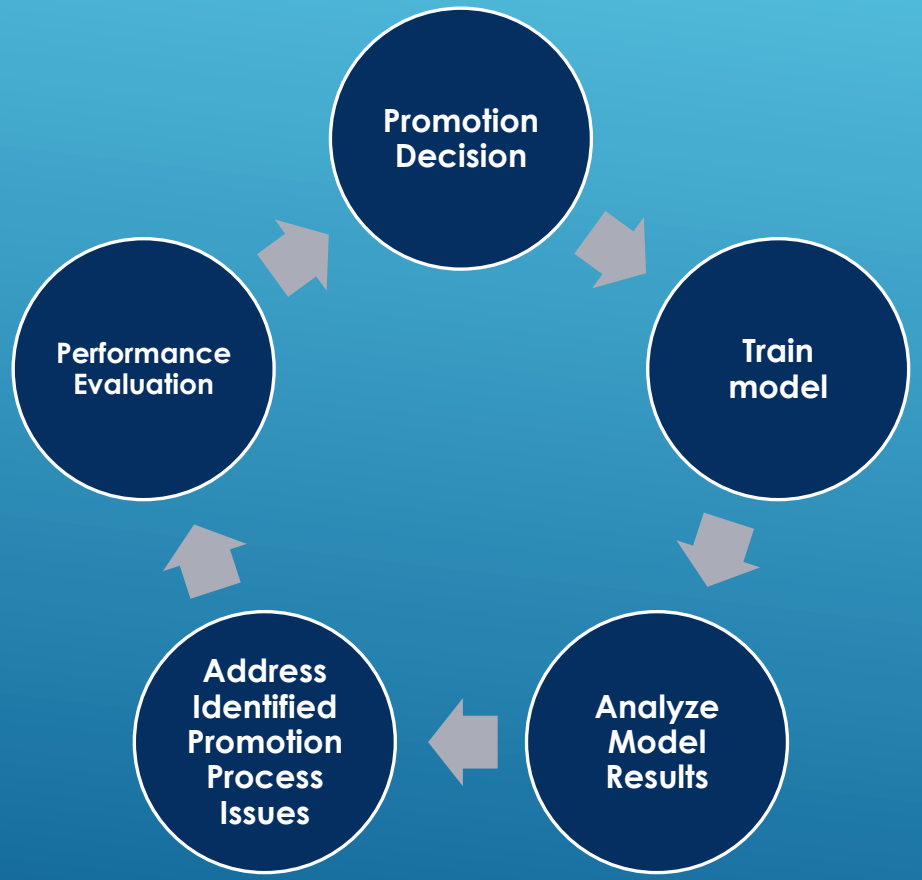


Source: https://www.paycor.com/resource-center/articles/mastering-9-box-method-key-to-effective-employee-coaching/

Organizational Process Recommendations

OPERATIONALIZING THE MODEL RESULTS
MODEL / PROMOTION PROCESS HARMONIZATION

Objective: Match actual promotion factors to model variables



OPERATIONALIZING THE MODEL RESULTS
PROPOSED ACTIONS

Decide:

1. Commit to an official organizational definition of merit
2. Establish or leverage existing metrics to represent the data definition of merit

Investigate:

1. Why don't merit variables carry more weight in promotions?
2. What other factors are being considered in promotions that are not currently accounted for?
3. Why are employees being penalized for recognized meritorious achievements?

Monitor:

1. Re-train the model each performance cycle with new data
2. Review model performance metrics - maintain 89% recall, increase Pseudo R^2
3. Analyze and address promotion process deviations and inconsistencies

OPERATIONALIZING THE MODEL RESULTS
EXPECTED BENEFITS

- Increased accuracy, integrity, and defensibility of promotion decisions.
- Decreased legal liability through proactive compliance, avoiding regulatory penalties
- A consistent, reproducible, and transparent framework for talent classification
- Reduced attrition of high potential employees who were otherwise overlooked by the existing promotion process

Thank You!