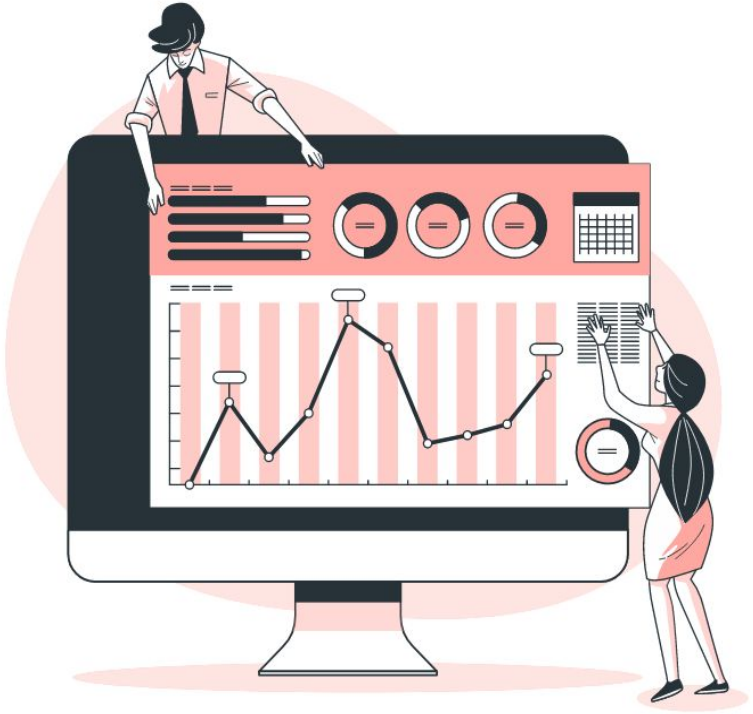
Employee Promotion Prediction

By Erlando Febrian



About Me

I graduated of bachelor's degree from Bandung Institute of Technology, School of Business and Management, Business degree. I also graduated from Rakamin Data Science bootcamp with outstanding grade, awarded as best final project team, and also my role as team leader. I experienced in the following scope:

- Supervised & Unsupervised Learning
- Time Series Forecasting
- A/B Testing
- Deep learning using TensorFlow and Pytorch
- Recommender System
- Customer Lifetime Value
- SQL & Data Visualization (Tableau & Power BI)

Contact


github.com/erlndofebri


brian-insights.site/


erlandoregita99@gmail.com


0821-1000-4094


linkedin.com/in/erlandoregita/

Table of Content

01

Background

Background

A **high number of promotions** are an indication that our company was **experiencing a lot of change**. **Acquiring new business or merging with another organization all require major shifts in labor allocation.** People will take on new tasks and be promoted into various roles based on new needs.[1]

Why a high promotion rate is important?

A high promotion rate means the company uses internal hiring so that employees can be promoted. Benefit of high promotion rate:

- Reduced training and socializing time.[2]
- Less external hiring cost.[3]
- Less probability to quit / be fired.[4]

1 Goals

Increase employees promotion rate up to **25%** and reduce total hiring cost up to **50%**

2 Metrics

- Promotion Rate
- Total External Hiring Cost

3 Objective

- Analyze factors that affect the increment of promotion rate
- Predict employees will be promoted or not using prediction model

Retaining Talent Is More Cost Effective Than Hiring [5]

- Rule of Thumb -

02

Exploratory Data Analysis

Dataset Overview

1 Year Historical Data contain of 54808 Employees (Rows)

Categorical Features

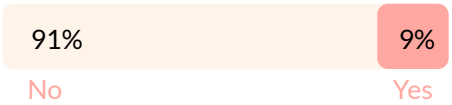
- Department
- Region
- Education
- Gender
- Recruitment Channel

Numerical Features

- Employee ID
- No of Training
- Age
- Previous Year Rating
- Length of Service
- Awards Won
- Avg Training Score

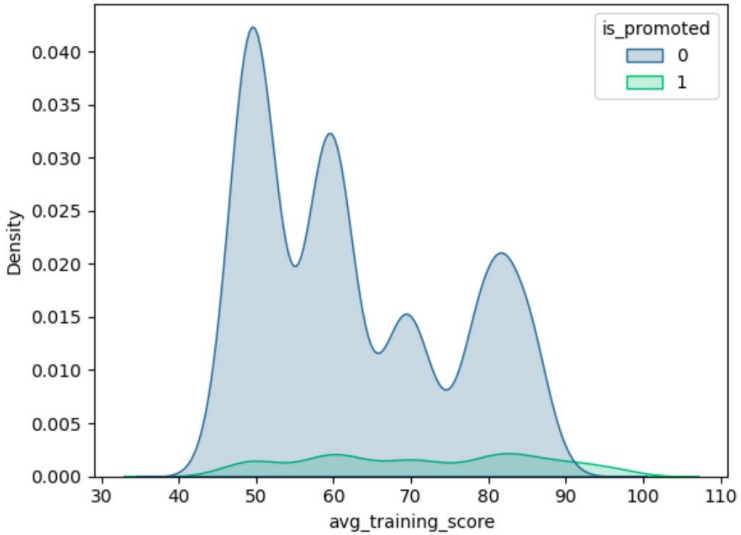
Target Feature

Is Promoted?



*) Detail Features Dictionary Written On Appendix

Data Exploration (Average Training Score)



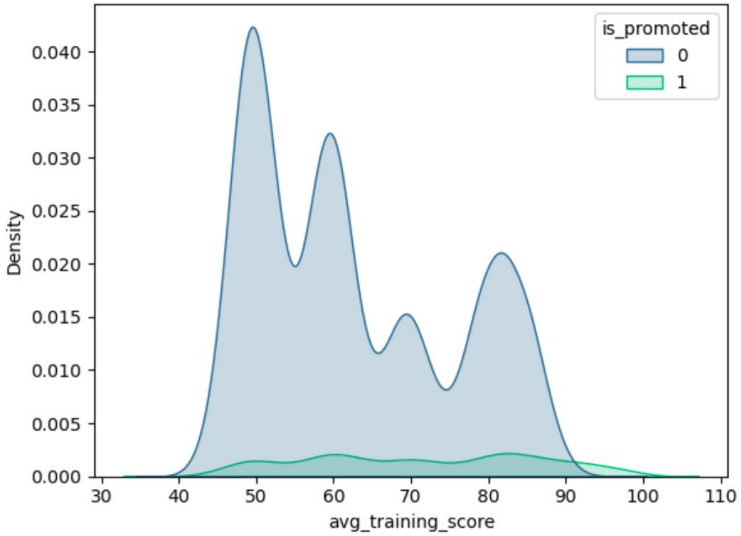
Employees who get promotion have higher avg_training_score, and known that the mean avg_training_score for Employees who get promotion are **71** and those who didn't get promotion are **62.**

Insights:

- Need to find the common causes why employees have low Average Training Score
- Provide several best recommendations to cope with low Average Training Score

**) Average Training Score: average score in current training evaluations*

Data Exploration (Previous Year Rating)



Employees who get promotion tend to have higher previous_year_rating, and known that the median previous_year_rating for Employees who get promotion are **4** and those who didn't get promotion are **3**

Insights:

- Need to find the common causes why employees have low Previous Year Rating
- Provide several best recommendations to cope with low Previous Year Rating

**) Previous Year Rating: Employee Rating for the previous year*

03

Data Pre Processing

Data Pre Processing

1

Handling missing Values

Using median and mode to do values imputation

2

Outlier Handling

We found that there are no extreme values and strage values. **We don't need to handle outlier values.**

3

Feature Encoding

Using One Hot Encoder: As a result, there are **51 additional encoded features**

Pre Processing

4 Feature Selection

Feature selection methods:

- **Filter method**
- **Embedded method**

5 Handling Imbalanced Target

Using **SMOTE with default sampling strategy** (1:1)

04

Modeling

Basic Modeling

We are using **2 experiments for basic modeling**, **first experiment is using all features (58 features)** and **second experiment using filtered features** from feature selection method

	Model	Accuracy	Precision	Recall	F1 Score
0	Logistic Regression	0.831296	0.208259	0.321233	0.252694
1	Decision Tree	0.887247	0.378844	0.421918	0.399222
2	Random Forest	0.935839	0.914110	0.306164	0.458697
3	Ada Boost	0.831296	0.208259	0.321233	0.252694
4	Gradient Boost	0.938089	0.943775	0.321918	0.480082
5	XG Boost	0.940826	0.907873	0.371233	0.526981

*) 1st experiment result (using all feature prediction)

Focus on **Precision** Score to minimize False Positive (Predicted **will promoted**, but **actually not promoted**). We consider **unqualified talent will get promotion**, and if it happen, **it will gain more cost** in the future. That's why we will minimize this type of employees. After that we will implement hyperparameter tuning on:

- Random Forest
- Gradient Boost
- XGBoost

Because they have high Precision Score

Hyperparameter Tuning

We are using **3 best model** from basic modeling

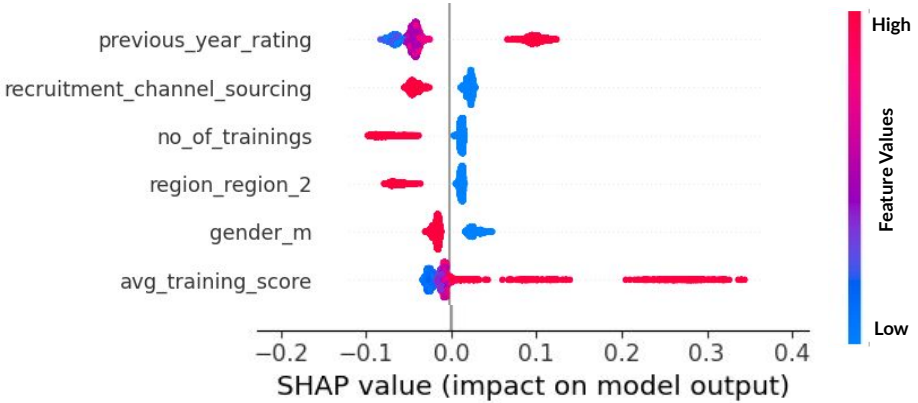
	Train Precision Score	Test Precision Score
Random Forest	**0.99**	**0.99**
Gradient Boost	0.85	0.83
XGBoost	0.96	0.97

After doing hyperparameter tuning, we decided to use **Random Forest** to be interpreted using SHAP values in the next step. The reason are :

- Best fit on precision train and test score
- Not overfitting / underfitting

Feature Importance

Top 5 Important Features



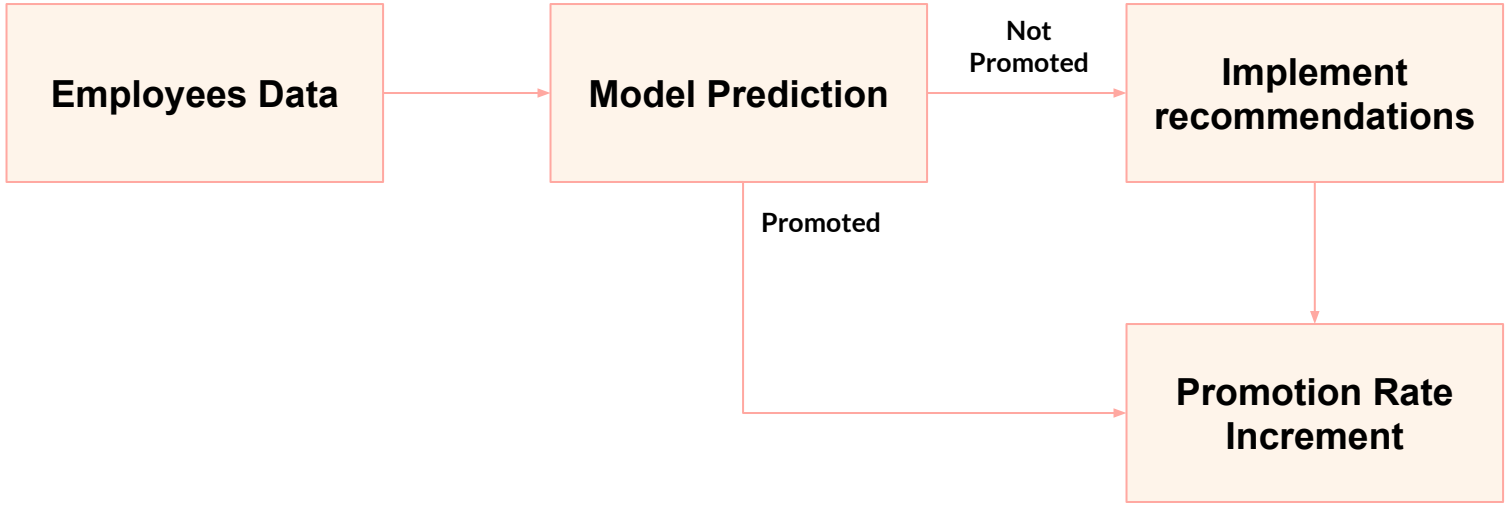
We can see that:

- Potential promotee who have high avg training score tend to be more promoted
- Potential promotee who have high previous year rating tend to be more promoted
- etc

05

Business Insights and Recommendation

How Our Model Works?



*) High promotion rate indicates the company is keep growing

Why you should use our Model & Recommendation?

These are the benefit comparison using and not using our Model

Factor	Using Model	Not Using Model
Importance Variable	Company know the important factors that affect employee promotion. By knowing this, company can provide effective treatment so that the promotion rate remains high	Company provides random treatment whose effectiveness cannot be accurately measured without an experiment. Ineffective treatment costs a lot of time and money
Cost	Company can save external hiring cost and training cost for new talents. They only have to focus on allocating costs in providing the effective treatment for existing employee and give them promotion	The longer the employee promotion cycle, the higher the cost
Time	Company needs help in identifying the eligible candidates at a particular checkpoint so that they can expedite the entire promotion cycle. By using our model, companies can speed up the employee promotion cycle. Employees who are detected not to be promoted will be given treatment as early as possible, so that in the near future the employee can be promoted (without external recruiting)	The final promotions are only announced after the evaluation and this leads to delay in transition to new roles

Root Cause Analysis

Low Promotion Rate

Why 1
Low Avg Training Score

Cause 1
Bad Training Systems [6]

Cause 2
Hard to get Access

Cause 3
Lack of leadership & soft skill training

Cause 4
Poor Training Content [7]

Why 2
Low Rating Previous Year

Cause 1
Not transparent performance review rating [8]

Cause 2
Subjective rating performance review [9]

Recommendation 1

Improve Average Training Score

Problem	Recommendation	Pros	Cons
Poor training system	Build and subscribe corporate LMS (Learning Management System)[10]	• Easy to track employees learning progress [11] • Unlimited learning access	• It cost a lot of money
Hard to get access	Always go mobile, make sure that our corporate LMS is mobile friendly and also make sure user friendly	• Employees can access the courses anytime anywhere • Increase probability of employee finishing the course	• It cost a time and money to build good LMS • Need additional time to do interface testing and gain LMS's content feedback

Recommendation 1

Improve Average Training Score

Problem	Recommendation	Pros	Cons
Lack of leadership and soft skill training	Create leadership and soft skill training course	• Improve employees performance • Employees with good leadership and soft skill tend to be more productive	• It cost time to build / arrange the course • Challenges to ensure employees complete leadership and soft skills training
Poor training content (Irrelevant training content)	Always find the relevant training course for employees. We can do some surveys to make sure it will be relevant for employees [12]	• Employees enjoy the learning process • Employees will get more benefit from the course / training	• It's hard and time consuming to find very relevant content for employees (because there is a time and cost limit)

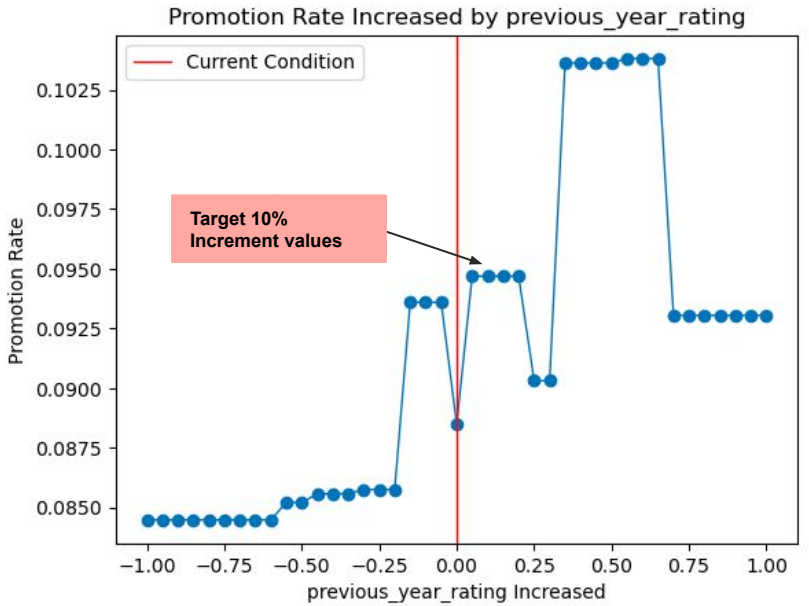
Recommendation 2

Improve Employee Performance Rating

Problem	Recommendation	Pros	Cons
No transparency employee performance rating	• Improve how performance feedback given • Constructive feedback • Be an active listener [12]	• Employee know what positif and negative side about them and know why they should improve it	• Being 100% transparent sometimes brings another problem
Subjective performance review rating	• Asking the right question • Create objective performance review system [13]	• Employee receive objective fact about them and know how to improve it	• It's hard to implement 100% objective for some type of employees

Sensitivity Analysis On Previous Year Rating

Previous Year Rating

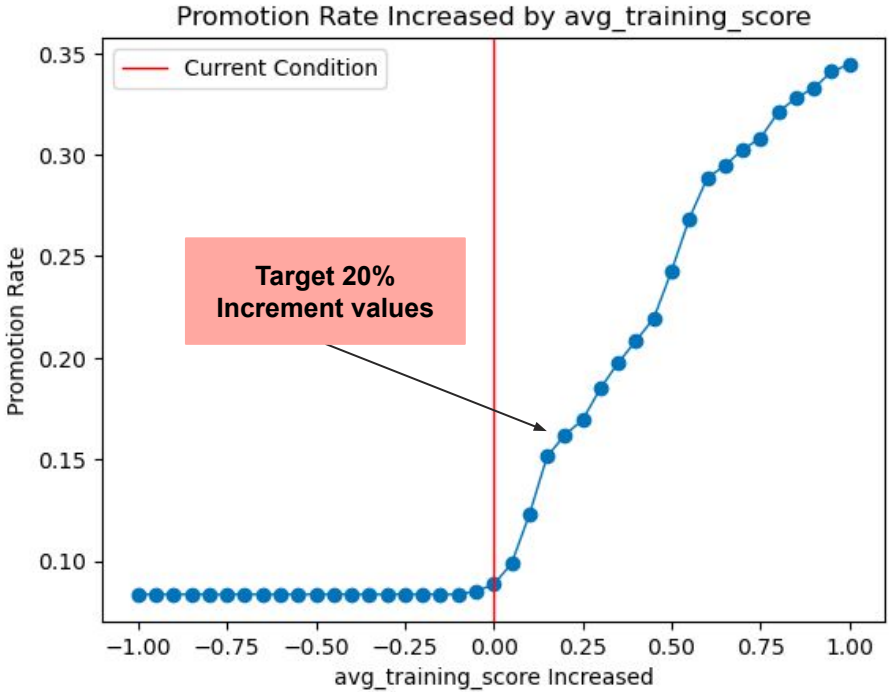


The **higher previous_year_rating** percentage increment, the **higher Promotion Rate** but **at some point** it's not that effective as before, then it will **reduce Promotion Rate**

Let's say we implement recommendations trough **previous_year_rating treatment**, we succeed **increase previous_year_rating by 10%,** then it affected to **Promotion Rate increment up to 1%.**

Sensitivity Analysis : Avg Training Score

Avg Training Score



The **higher avg_training_score** percentage increment, the **higher Promotion Rate** at the same time

Let's say we implement recommendations trough **avg_training_score treatment**, we succeed **increase avg_training_score by 20%,** then it affected to **Promotion Rate increment up to 8.12%.** And so on..

Metrics and Assumption on Simulation

Before running the simulation, here the assumption and metrics we need to know

Assumption

1 Average Cost Per Hiring

$ 4,425 / Hire [14]

2 External Hiring Target Percentage

Max 2% of Remaining Employees Who Didn't Get Promotion

3 Total Num of Employees

54808 Employees (Gained from total rows of the data)

Metrics

4 Num of Recruited Talent

Maximum number of talents that can be recruited.

5 Total External Hiring Cost

Total cost needed to recruit all talent

6 Promotion Rate

Percentage of employees who get promoted over all the employees

*) **4 Formula** : **num_recruited_talent** = hiring_target_pctg * (1-promotion_rate) * num_of_employees
*) **5 Formula: total_hiring_cost** = num_recruited_talent * avg_cost_per_hiring

First Business Simulation

Improve 10% Increment on Average Training Score

BEFORE		***AFTER***
8.52% Promotion Rate / Year	➡	**12.41%** Promotion Rate / Year
44M Hiring Cost / Year	➡	**42M** Hiring Cost / Year

Implementation :

1. *Current avg_training_score* mean : 60
2. *Avg_training_score* mean after increasing 10% average : 70
3. Current *num_recruited_talent* : 10,028 talents
4. *Num_recruited_talent* after increasing promotion rate: 9,601 talents

With **10% improvement on avg_training_score,** it potentially increase **3.9% Promotion Rate** and decrease **4.45 % Hiring Cost**.

Second Business Simulation

Improve 20% Increment on Average Training Score

BEFORE		AFTER
8.52% Promotion Rate / Year	➡	**16.64%** Promotion Rate / Year
44M Hiring Cost / Year	➡	**40M** Hiring Cost / Year

Implementation :

1. *Current avg_training_score* mean : 60
2. *Avg_training_score* mean after increasing 20% average : 76
3. Current *num_recruited_talent* : 10,028 talents
4. *Num_recruited_talent* after increasing promotion rate: 9,138 talents

With **20% improvement on avg_training_score,** it potentially increase **8.12% Promotion Rate** and decrease **9.74 % Hiring Cost**.

Third Business Simulation

Improve 30% Increment on Average Training Score

BEFORE		***AFTER***
8.52% Promotion Rate / Year	➡	**18.95%** Promotion Rate / Year
44M Hiring Cost / Year	➡	**39M** Hiring Cost / Year

Implementation :

1. *Current avg_training_score* mean : 60
2. *Avg_training_score* mean after increasing 30% average : 82
3. Current *num_recruited_talent* : 10,028 talents
4. *Num_recruited_talent* after increasing promotion rate: 8,884 talents

With **30% improvement on avg_training_score,** it potentially increase **10.43% Promotion Rate** and decrease **12.87 % Hiring Cost**.

Fourth Business Simulation

Improve 10% Increment on Previous Year Rating

BEFORE		***AFTER***
8.52% Promotion Rate / Year	➡	**9.52%** Promotion Rate / Year
44M Hiring Cost / Year	➡	**43M** Hiring Cost / Year

Implementation :

1. *Current previous_year_rating* mean : 3
2. *previous_year_rating* mean after increasing 10% average : 4
3. Current *num_recruited_talent* : 10,028 talents
4. *Num_recruited_talent* after increasing promotion rate: 9,918 talents

With **10% improvement on previous_year_rating,** it potentially increase **1% Promotion Rate** and decrease **1.11 % Hiring Cost**.

Fifth Business Simulation (Combined)

Improve 10% Increment on Previous Year Rating and 20% Increment on Average Training Score

BEFORE		***AFTER***
8.52% Promotion Rate / Year	➡	**21.30%** Promotion Rate / Year
44M Hiring Cost / Year	➡	**38M** Hiring Cost / Year

Implementation :

1. *previous_year_rating* mean after increasing 10% average : 4
2. *Avg_training_score* mean after increasing 13.5% average : 72
3. *Num_recruited_talent* after increasing promotion rate: 8,627 talents

With these improvement, it potentially increase **12.78% Promotion Rate** and decrease **16.24% Hiring Cost**.

Sixth Business Simulation (Combined)

Improve 10% Increment on Previous Year Rating and 30% Increment on Average Training Score

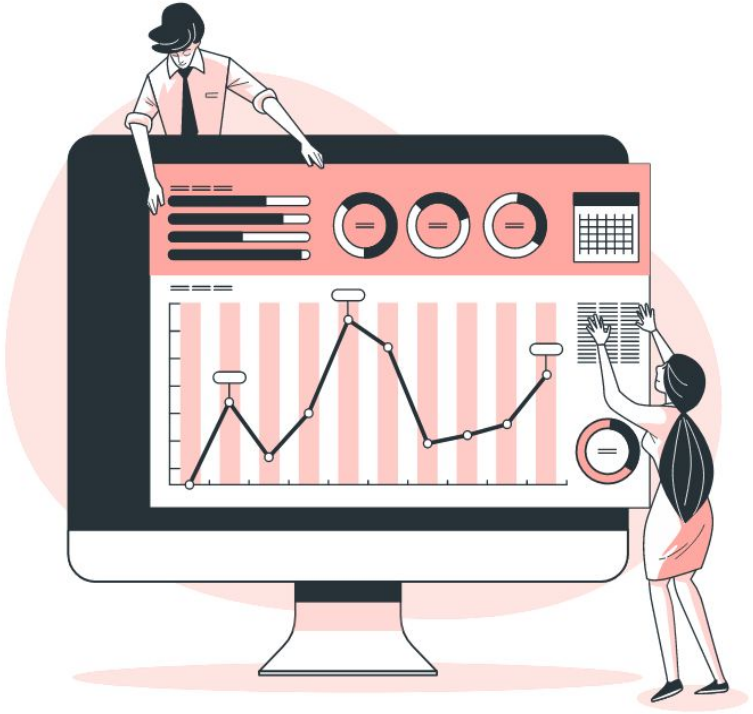
BEFORE		***AFTER***
8.52% Promotion Rate / Year	➡	**24.63%** Promotion Rate / Year
44M Hiring Cost / Year	➡	**36M** Hiring Cost / Year

Implementation :

1. *previous_year_rating* mean after increasing 10% average : 4
2. *Avg_training_score* mean after increasing 30% average : 82
3. *Num_recruited_talent* after increasing promotion rate: 8,262 talents

With these improvement, it potentially increase **16.11% Promotion Rate** and decrease **21.38 % Hiring Cost**.

THANK YOU



Appendix

Feature Dictionary

- employee_id: Unique ID for employee
- department: Department of employee
- region: Region of employment (unordered)
- education: Education Level
- gender: Gender of Employee
- recruitment_channel: Channel of recruitment for employee
- no_of_trainings: No of other trainings completed in previous year
- age: Age of Employee
- previous_year_rating: Employee Rating for the previous year
- length_of_service: Length of service in years
- awards_won?: If awards won during previous year then 1 else 0
- avg_training_score: Average score in current training evaluations
- is_promoted: (Target) Recommended for promotion

References

1. https://www.employeecycle.com/an_hr_metric_promotion_rate/
2. https://www.linkedin.com/pulse/20141029193238-26473352-external-recruiting-vs-internal-promotion/
3. https://employmenthero.com/blog/recruitment-costs/
4. https://hrdailyadvisor.blr.com/2019/07/01/pros-and-cons-to-internal-promotion-vs-external-hiring/
5. https://www.tlnt.com/retaining-talent-is-more-cost-effective-than-hiring/
6. https://elearningindustry.com/speed-up-skills-training-and-boost-employee-performance
7. https://www.westnetlearning.com/employee-training/
8. https://www.quantumworkplace.com/future-of-work/13-performance-review-tips-that-actually-improve-employee-performance
9. https://engagedly.com/employee-performance-review-tips-to-improve-performance/
10. https://elearningindustry.com/speed-up-skills-training-and-boost-employee-performance

References

11. https://whatfix.com/blog/corporate-learning-management-systems/
12. https://engagedly.com/employee-performance-review-tips-to-improve-performance/
13. https://www.quantumworkplace.com/future-of-work/13-performance-review-tips-that-actually-improve-employee-performance
14. https://www.zippia.com/advice/cost-of-hiring-statistics-average-cost-per-hire/