

# Probabilistic Turnover Risk Scoring with Tiered Classification

FINAL PROJECT REPORT – RAKAMIN DATA SCIENCE BATCH 55

July - August

Mentor : Dedi Irwanto & Rahmat Faizal

# MEMBER

TalentaHub



SUKMA INDAH

PM  
&  
Data  
Engineer



FITRIA SHINTA  
ARIZQI

Business  
Analyst



NAUFAL IFFA  
M.R

Data  
Engineer



FERRY  
IRWANTO

Data  
Science



Predict.Prevent.Retain

# OUTLINE

---

- Stage 0 : Project Initiation & Problem Solving
- Stage 1 : Data Acquisition & Preparation
- Stage 2 : Model Development & Experimentation
- Stage 3 : Model Evaluation & Interpretability
- Stage 4 : Deployment Evaluation & Interpretability
- THANK YOU

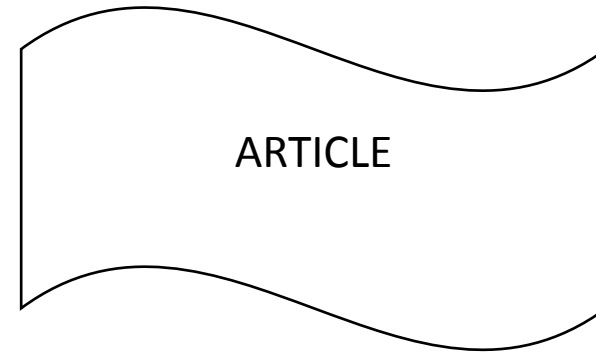
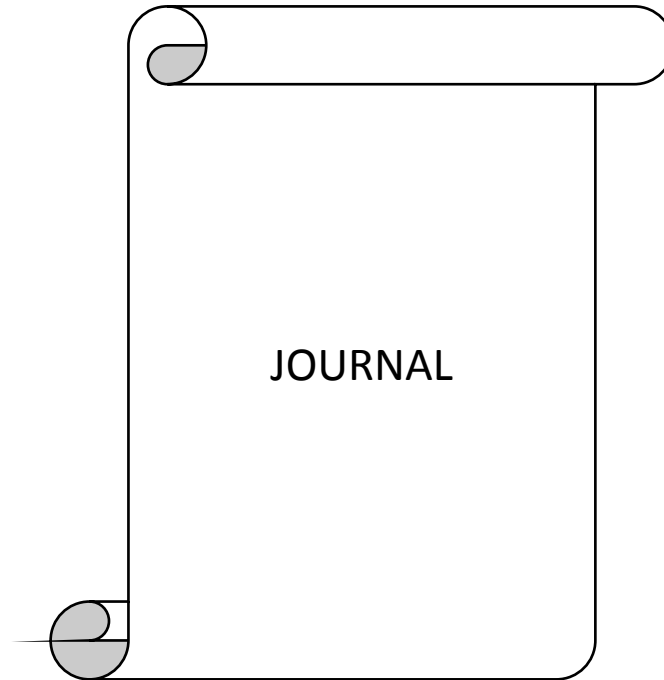
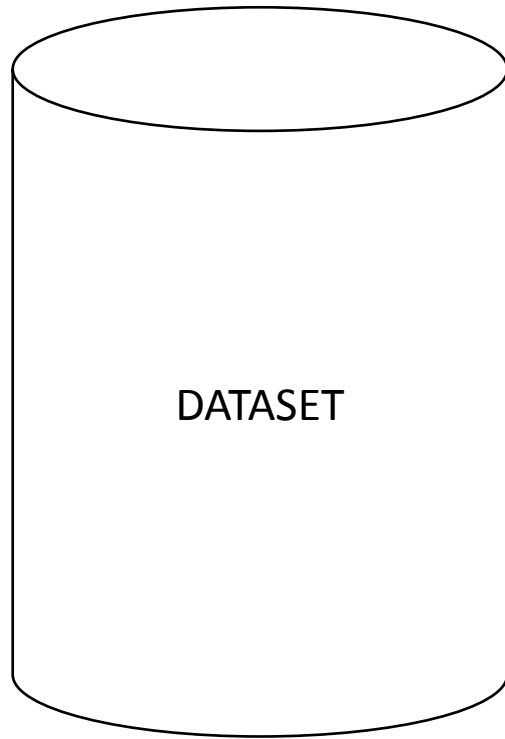
# Stage 0 : Project Initiation & Problem Solving



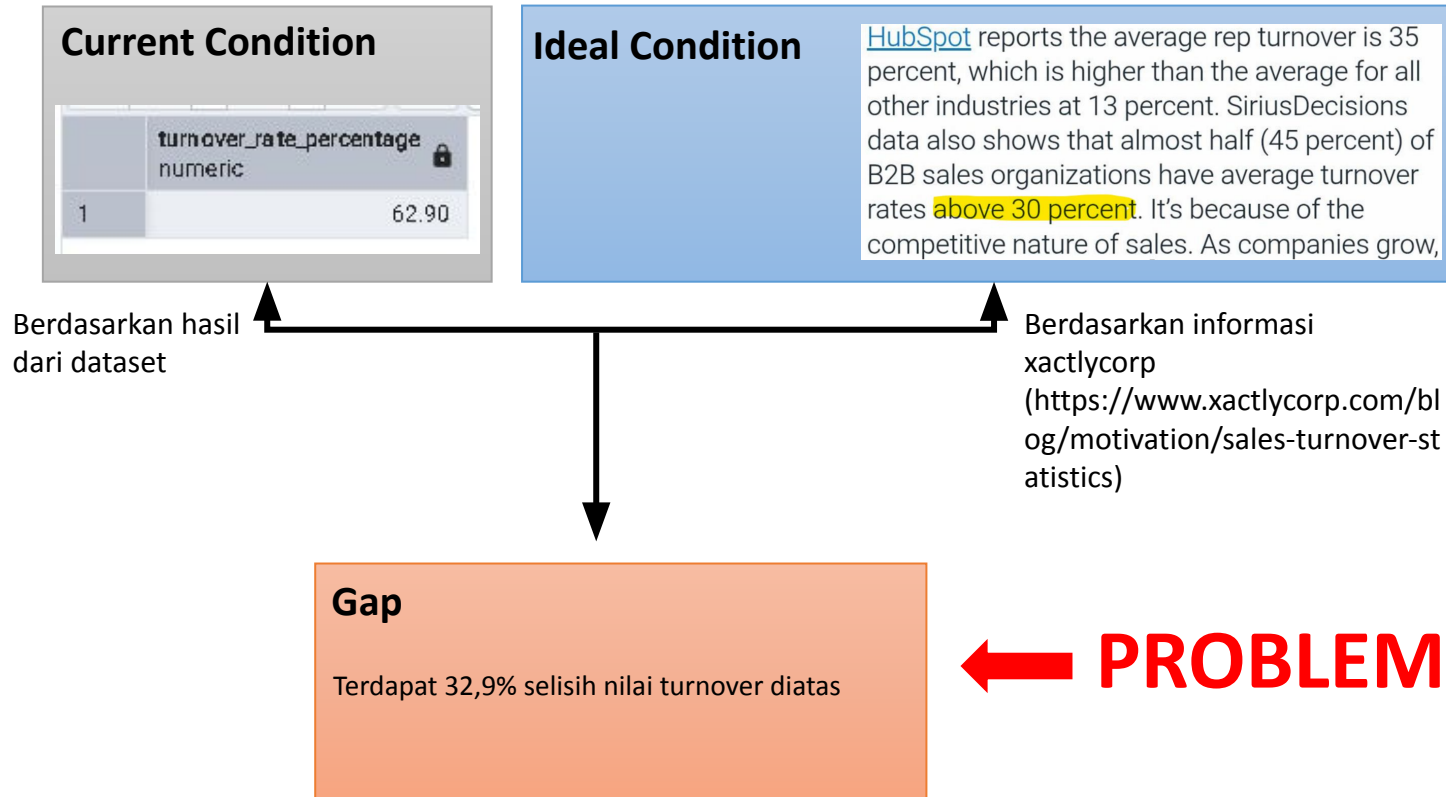
# RESEARCH

---

TalentaHub



# PROBLEM



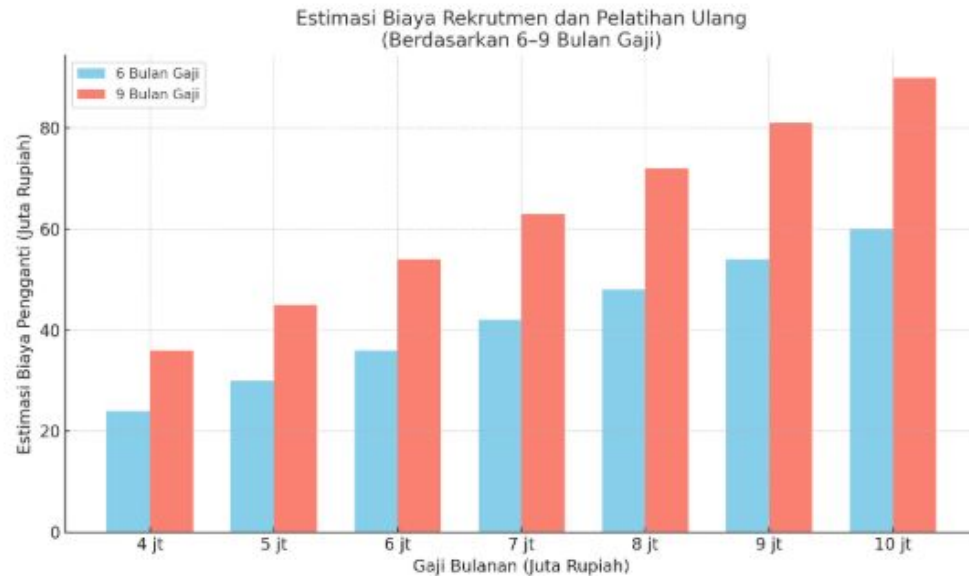
## Problem Statement

TalentaHub adalah perusahaan yang bergerak di bidang penyedia layanan digital untuk rekrutmen dan manajemen talenta khusus di bidang sales. Pada akhir tahun kemarin, perusahaan mengalami peningkatan terhadap **turnover rate 32,9%**. **Turnover** merupakan tingkat di mana karyawan meninggalkan sebuah perusahaan dan digantikan oleh karyawan baru (<https://sundayinsurance.co.id/blog/turnover-adalah>). Tingginya turnover rate berdampak pada meningkatnya biaya rekrutmen, pelatihan ulang, serta terganggunya kontinuitas dan produktivitas tim.

Meskipun TalentaHub telah menyediakan kompensasi kompetitif dan program internal engagement, belum ada pendekatan berbasis data yang digunakan untuk secara proaktif mengidentifikasi karyawan yang berisiko tinggi untuk resign.

# PROBLEM

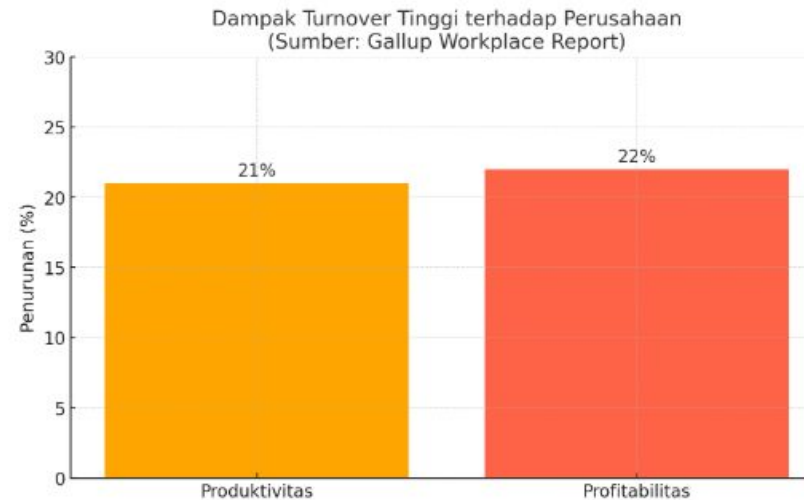
Dampak dari turnover karyawan :



## Biaya Rekrutmen & Pelatihan Ulang:

Rata-rata biaya untuk menggantikan karyawan adalah **6 hingga 9 bulan gaji** karyawan tersebut.

(Sumber: **Society for Human Resource Management (SHRM)**, 2022)



## Gangguan terhadap Produktivitas & Tim:

Perusahaan dengan tingkat turnover tinggi mengalami penurunan produktivitas hingga **21%**, dan penurunan profitabilitas hingga **22%**.

(Sumber: **Gallup Workplace Report**)

Sehingga **solusi** terbaik dari permasalahan ini adalah **mempertahankan karyawan**



# BUSINESS UNDERSTANDING

## GOALS

Untuk membangun sistem yang mampu melakukan **profiling risiko karyawan** berdasarkan kemungkinan mereka untuk mengundurkan diri (resign) dari perusahaan. Dengan adanya profil risiko ini, kami dapat:

- Mengidentifikasi karyawan yang memiliki potensi tinggi untuk resign.
- Menganalisis faktor-faktor yang mempengaruhi keputusan karyawan untuk keluar.
- Meningkatkan strategi retensi

## OBJECTIVE

Untuk mencapai goals tersebut, kami memiliki beberapa sasaran utama, yaitu:

- Membangun **model machine learning** yang dapat memprediksi probabilitas karyawan akan resign berdasarkan data historis karyawan
- Melakukan **segmentasi risiko** untuk mengklasifikasikan karyawan ke dalam beberapa level risiko (misalnya: low, medium, high).
- Menyediakan **profil risiko karyawan** yang dapat dijadikan dasar dalam pengambilan keputusan strategis oleh manajemen HR.
- Memberikan **dashboard atau laporan visualisasi** yang mudah dipahami guna membantu tim HR memahami hasil prediksi dan segmentasi risiko.

## METRIC BUSINESS

- Alokasi cost dengan tepat sasaran , sesuai dengan prinsip pareto dimana 20 % turnover karyawan dapat menyebabkan 80% kerugian perusahaan, maka dari itu pengalokasi budgeting untuk pelatihan kepada segmentasi turnovernya :

High	– 60%
Medium	– 30%
Low	– 10%

- Menurunkan turnover rate hingga dibawah 30%

# BUSINESS UNDERSTANDING

Prediksi Untuk	Time Window (Input)	Snapshot Date	Prediction Window
Jan-25	Oct-Dec 2024	31-Dec-24	Jan-25
Feb-25	Nov-Jan	31-Jan-25	Feb-25
Mar-25	Dec-Feb	28-Feb-25	Mar-25
Apr-25	Jan-Mar	31-Mar-25	Apr-25
Mei 2025	Feb-Apr	30-Apr-25	Mei 2025
Jun-25	Mar-May	31-May-25	Jun-25
Jul-25	Apr-Jun	30-Jun-25	Jul-25
Ags 2025	Mei-Jul	31-Jul-25	Aug-25
dst	dst	dst	dst

Time window : inputan data 3 bulan

Snapshot : tanggal akhir inputan data

Prediction window: waktu(bulan) prediksi resign

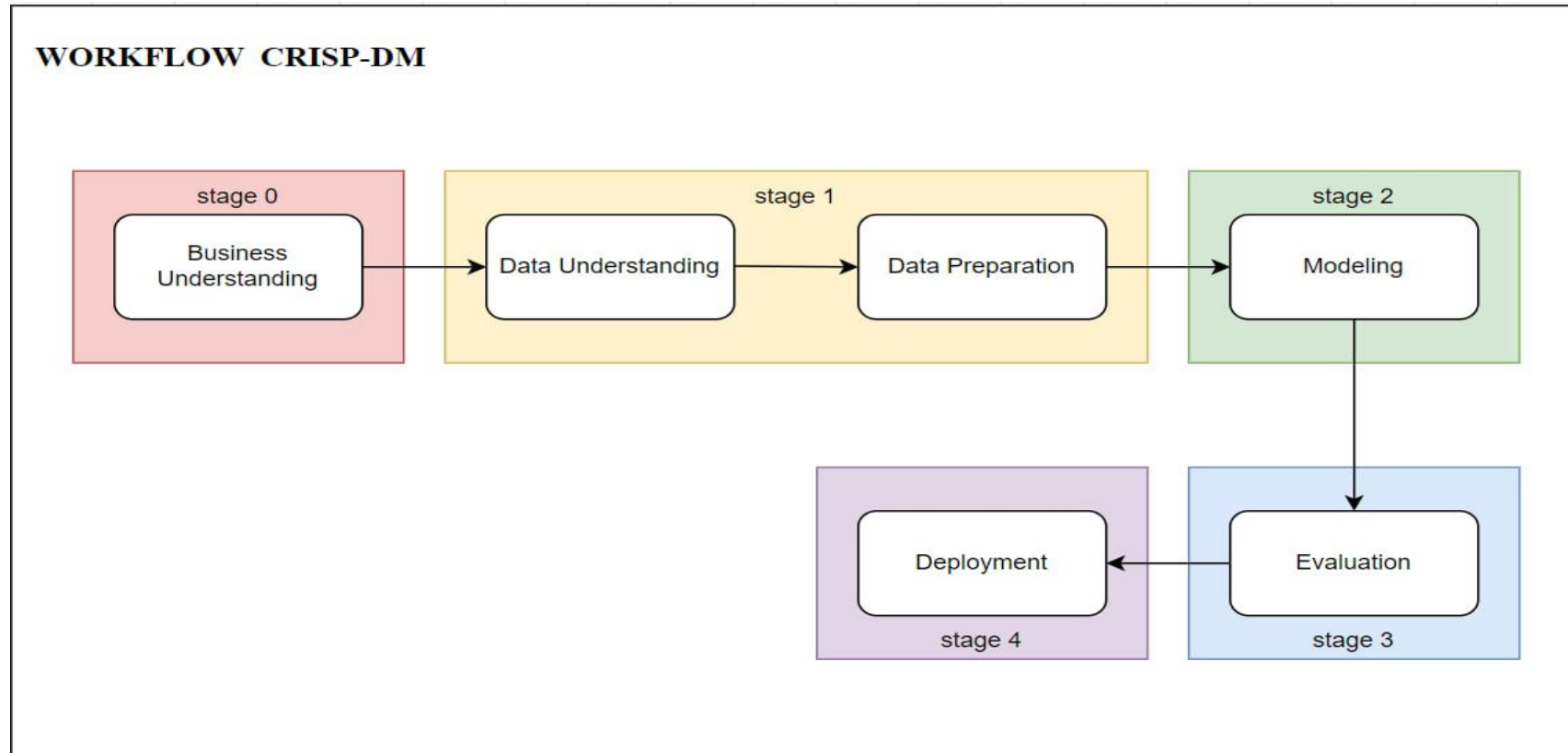
# TIMELINE

No	Nama Aktivitas	Nama Task	Role	start date	due date	PIC	Progress Task
1	Stage 0 : Project Initiation & Problem Framing	Riset Dataset	All Role	6/30/2025	7/1/2025	Sukma	Done
		Explorasi Dataset Untuk Konteks Bisnis	Bussiness Analyst	6/30/2025	7/1/2025	Shinta	Done
		Merumuskan Tujuan Bisnis (Business Objective)	Bussiness Analyst	6/30/2025	7/1/2025	Adit	Done
		Identifikasi Masalah yang Akan Diselesaikan dengan Data Science	Data Scientist	7/1/2025	7/2/2025	Ferry	Done
		Identifikasi Masalah yang Akan Diselesaikan dengan Data Science	Data Engineer	7/1/2025	7/2/2025	Naufal	Done
		Menentukan Metric Keberhasilan	All Role	7/1/2025	7/2/2025	Shinta & Adit	Done
		Membuat Laporan dan PPT	Project Manager	7/2/2025	7/3/2025	Sukma	In Prog...
		Validasi Awal dengan Mentor	All Role	7/4/2025	7/4/2025	All Member	Not Yet
		Memperbaiki sesuai feedback	All Role	7/5/2025	7/6/2025	All Member	Not Yet
2	Stage 1: Data Acquisition & Preparation	Data wrangling (handling missing values, duplicates, outliers).	Data Engineer	7/7/2025	7/8/2025	Sukma	Not Yet
		Feature engineering (one-hot encoding, feature scaling, feature extraction).	Data Engineer	7/7/2025	7/8/2025	Naufal	Not Yet
		Data visualization untuk memahami distribusi dan trend dalam data.	Data Scientist	7/8/2025	7/9/2025	Ferry	Not Yet
		Exploratory Data Analysis (EDA) report.	Bussiness Analyst	7/9/2025	7/10/2025	Shinta & Adit	Not Yet
		Mentoring Session	All Role	7/11/2025	7/11/2025	All Member	Not Yet
		Memperbaiki sesuai feedback	All Role	7/12/2025	7/13/2025	All Member	Not Yet

MANPOWER

TIMELINE

# WORKFLOW



# PROJECT RISK

No	Kategori	Risiko	Dampak	Mitigasi
1	Data	<ul style="list-style-type: none"><li>- Data Quality</li><li>- Data Drive</li></ul>	<ul style="list-style-type: none"><li>- Model tidak dapat belajar dengan baik</li><li>- Data bisa menyesatkan jika dibaca tanpa konteks atau pemahaman analitis yang cukup</li></ul>	<ul style="list-style-type: none"><li>- Preprocess data dengan benar serta pelajari fungsi fitur yang tersedia</li><li>- Selalu uji hasil analisis dengan realita lapangan</li></ul>
2	Model	<ul style="list-style-type: none"><li>- Overfit / Underfit</li><li>- Bias</li></ul>	Hasil evaluasi tidak sesuai	Melakukan testing dan modifikasi model
3	User	<ul style="list-style-type: none"><li>- Ditolak</li></ul>	Project tidak berjalan	Sediakan data dan bukti pendukung yang sesuai dengan kebutuhan user

# FEASIBILITY ANALYSIS

## 1. Kelayakan Teknis :

Dataset : tersedia dengan variable yang memadai

Tools : google colab

## 2. Kelayakan Manpower :

Terdapat 5 orang manpower

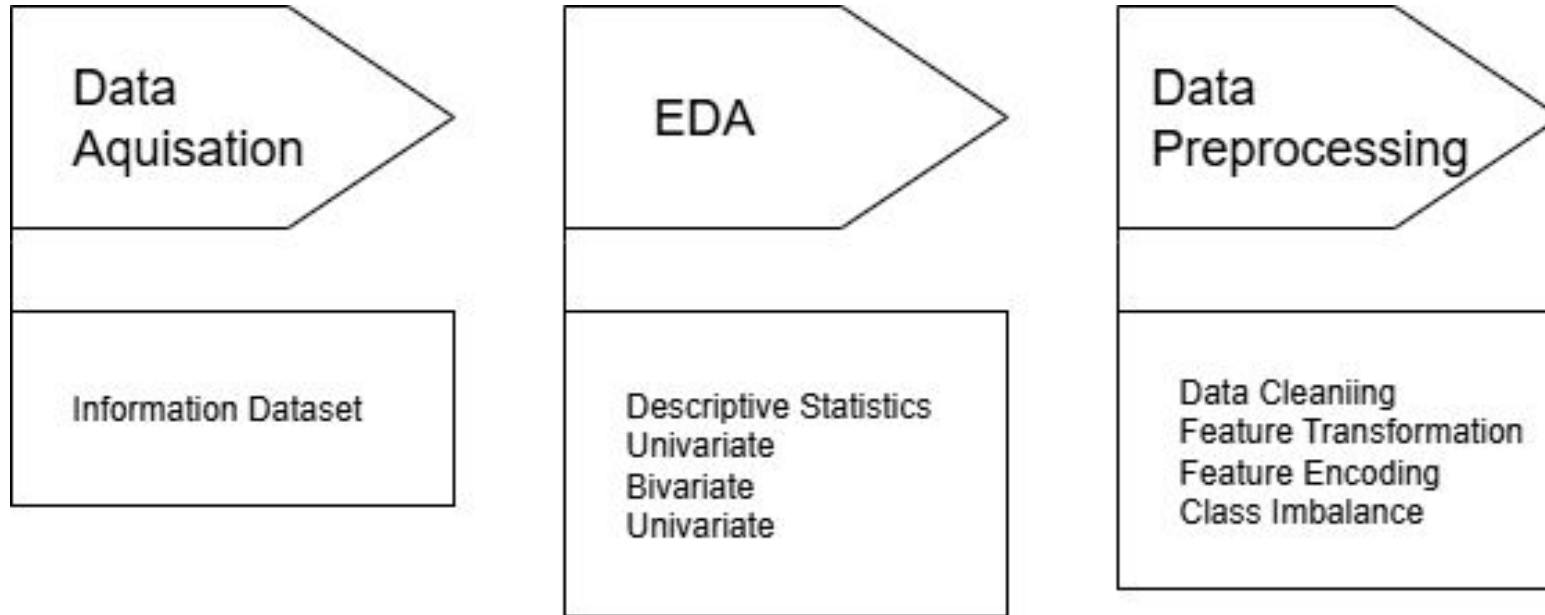
## 3. Kelayakan Waktu:

Waktu penyelesaian project selama 5 minggu

Berdasarkan informasi diatas, project ini **layak** untuk dijalankan

# Stage 1 : Data Acquisition & Preparation

# STEP





# INFORMATION DATASET

TalentaHub

Kolom	Tipe Data	Deskripsi
employee_id	Int	ID unik untuk tiap karyawan. Digunakan untuk identifikasi individual.
age	Int	Umur karyawan dalam tahun. Bisa digunakan untuk melihat hubungan usia dengan churn.
gender	String	Jenis kelamin karyawan. Bisa jadi fitur demografis untuk analisis.
education	String	Tingkat pendidikan terakhir. Bisa memengaruhi ekspektasi kerja & loyalitas.
experience_years	Int	Lama pengalaman kerja sebelum masuk perusahaan saat ini. Indikator kematangan profesional.
monthly_target	Int	Target kerja yang harus dicapai per bulan (misal: sales, unit, dsb). Digunakan untuk evaluasi performa.
target_achievement	Float	Persentase pencapaian terhadap target bulanan. Bisa berkorelasi dengan kepuasan kerja atau churn.
working_hours_per_week	Int	Total jam kerja per minggu. Tingkat beban kerja bisa memengaruhi churn.
overtime_hours_per_week	Int	Total jam lembur per minggu. Terlalu banyak lembur bisa menjadi faktor stres dan churn.
salary	Int	Gaji pokok bulanan karyawan. Faktor finansial yang berpengaruh terhadap churn.
commission_rate	Float	Persentase komisi tambahan (misal untuk role sales). Semakin tinggi, bisa meningkatkan retensi.
job_satisfaction	Int	Skor kepuasan kerja dari survei internal. Biasanya berkorelasi kuat dengan churn.
work_location	String	Tipe lokasi kerja. Bisa memengaruhi fleksibilitas dan loyalitas.
manager_support_score	Int	Penilaian terhadap dukungan dari atasan. Dukungan manajerial yang buruk bisa tingkatkan churn.
company_tenure_years	Float	Lama kerja di perusahaan ini (dalam tahun). Bisa jadi indikator loyalitas.
churn	Int	Apakah karyawan tersebut keluar dari perusahaan (1 = churn, 0 = stay). Target utama analisis.
marital_status	String	Status pernikahan. Terkadang dikaitkan dengan kebutuhan fleksibilitas atau lokasi kerja.
distance_to_office_km	Int	Jarak dari rumah ke kantor. Jarak jauh bisa memicu churn terutama pada role onsite.
churn_period	String	Waktu kapan karyawan keluar (jika churn). Penting untuk klasifikasi churn awal vs lambat.

## Descriptive Statistics

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 19 columns):
#   Column                Non-Null Count  Dtype
---  -
0   employee_id            1000 non-null   int64
1   age                    1000 non-null   int64
2   gender                  1000 non-null   object
3   education               1000 non-null   object
4   experience_years        1000 non-null   int64
5   monthly_target         1000 non-null   int64
6   target_achievement     1000 non-null   float64
7   working_hours_per_week 1000 non-null   int64
8   overtime_hours_per_week 1000 non-null   int64
9   salary                  1000 non-null   int64
10  commission_rate        1000 non-null   float64
11  job_satisfaction        1000 non-null   int64
12  work_location           1000 non-null   object
13  manager_support_score   1000 non-null   int64
14  company_tenure_years    1000 non-null   float64
15  churn                   1000 non-null   int64
16  marital_status          1000 non-null   object
17  distance_to_office_km   1000 non-null   int64
18  churn_period            1000 non-null   object
dtypes: float64(3), int64(11), object(5)
memory usage: 148.6+ KB
```

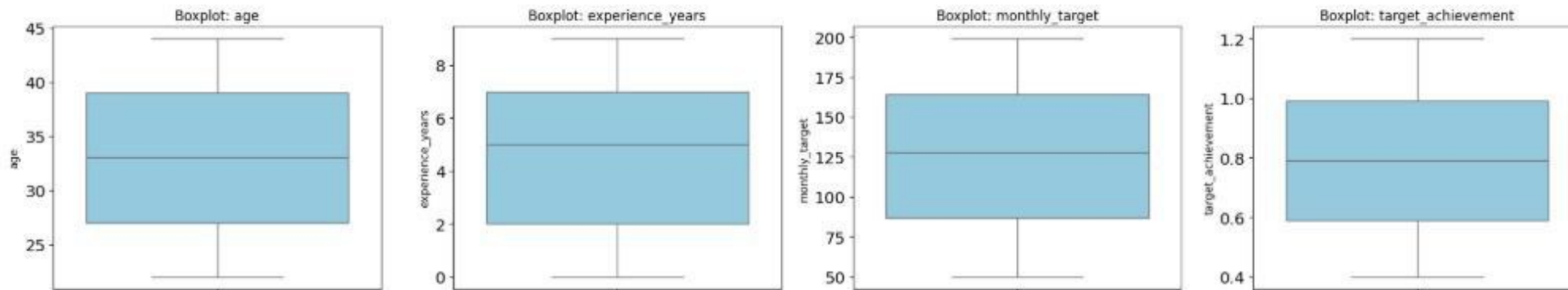
Dari informasi yang diberikan didapati bahwa data **tidak** ada yang **null** dan **tipe data** yang dipakai sudah **sesuai** serta **tidak** terdapat data yang **duplikasi**, sehingga tidak diperlukan untuk menambah data jika null atau mengganti tipe data serta menghapus data duplikat

```
✓ [15] data.duplicated().any()
0: np.False_
```

## Descriptive Statistics

```
employee_id: 0 outlier
age: 0 outlier
experience_years: 0 outlier
monthly_target: 0 outlier
target_achievement: 0 outlier
working_hours_per_week: 0 outlier
overtime_hours_per_week: 0 outlier
salary: 0 outlier
commission_rate: 0 outlier
job_satisfaction: 0 outlier
manager_support_score: 0 outlier
company_tenure_years: 0 outlier
churn: 0 outlier
distance_to_office_km: 0 outlier
```

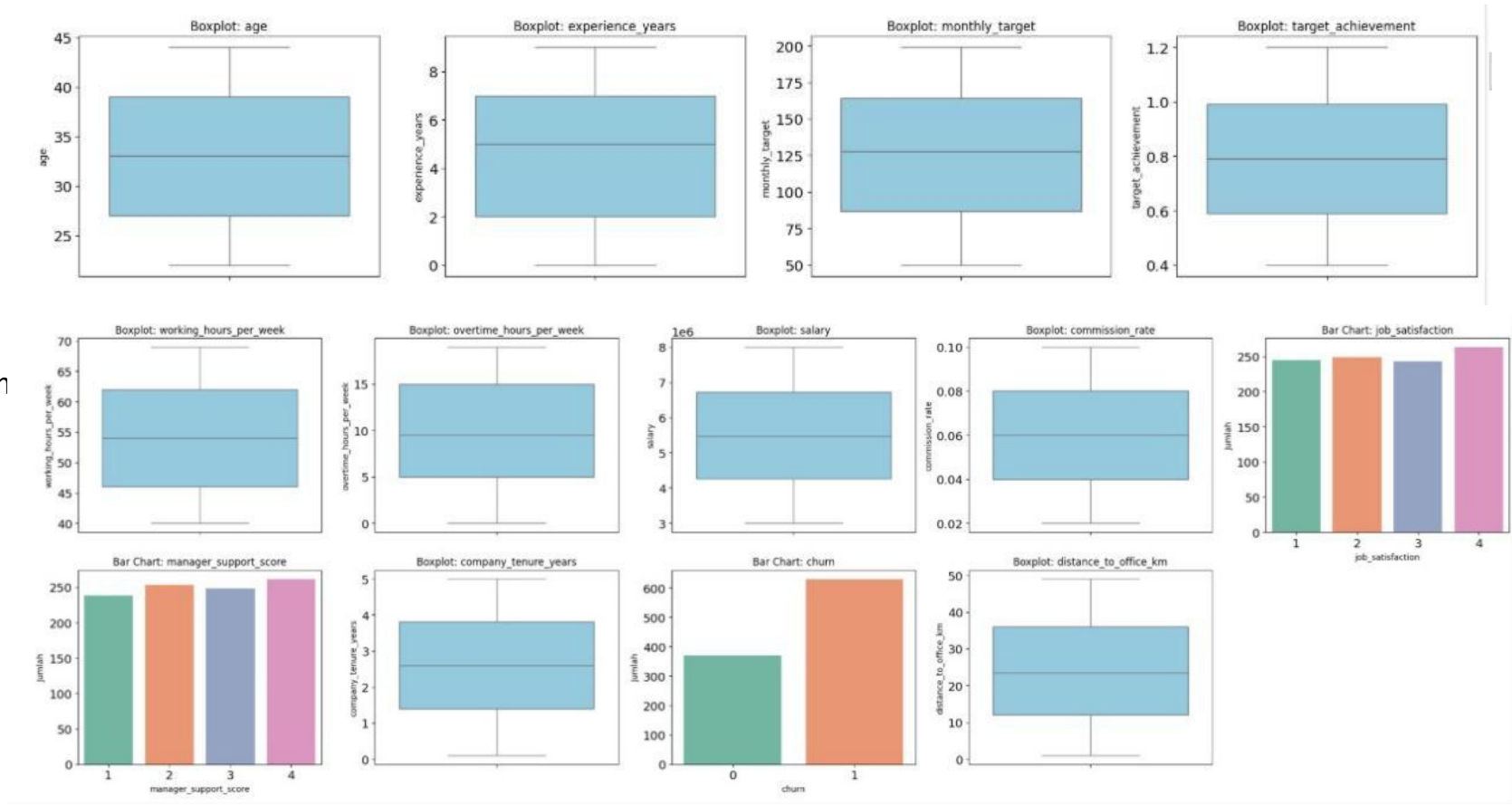
Dilakukan pengecekan terhadap data **outlier** dengan perhitungan dan grafik, di dapati bahwa data **tidak** ada data yang outlier dimiliki oleh dataset, sehingga tidak perlu dilakukan penghapusan data outlier.



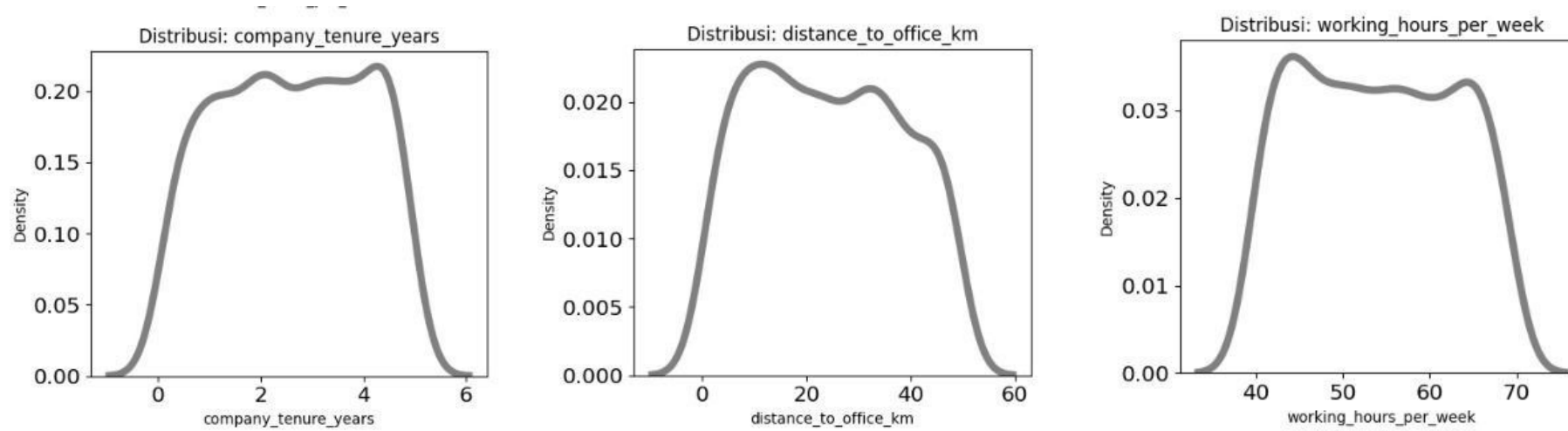
# EDA

## Univariate

Dari hasil grafik , dapat diketahui **rentang nilai** dalam **batas wajar**

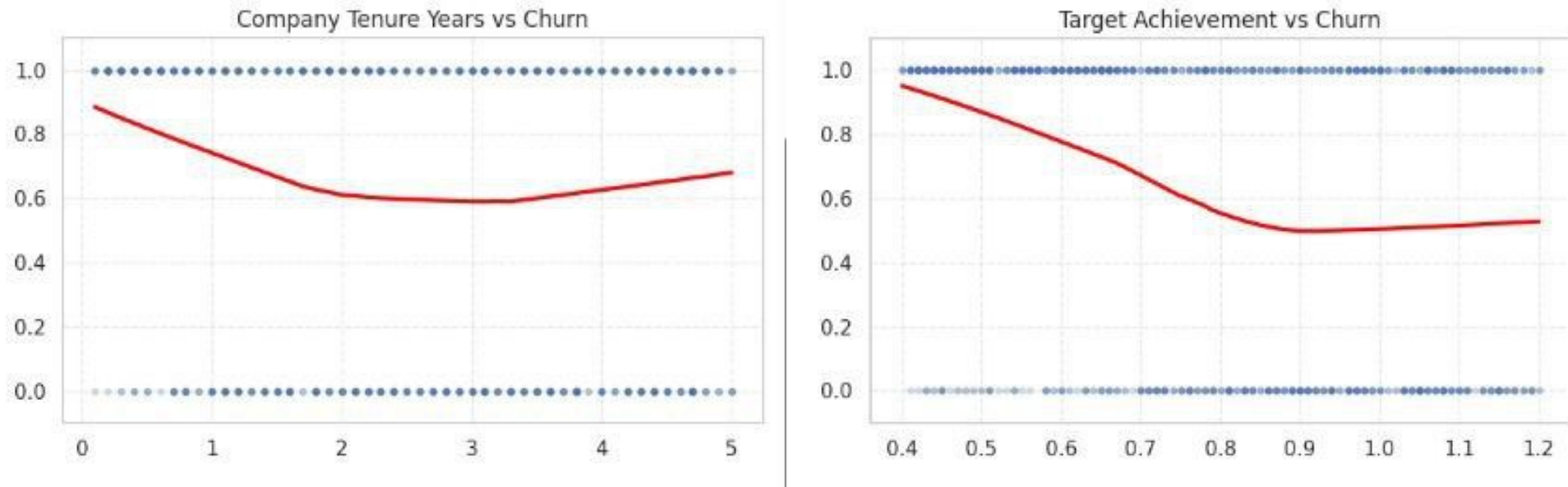


## Univariate



Ketika dilihat dari distribusi plot data termasuk **simetris**  
**mendekati normal**, sehingga **tidak** perlu dilakukan  
**Feature Transformation**

## Bivariate

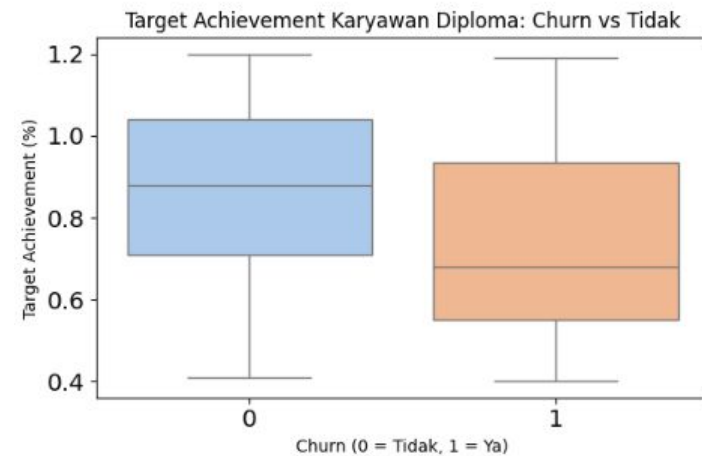
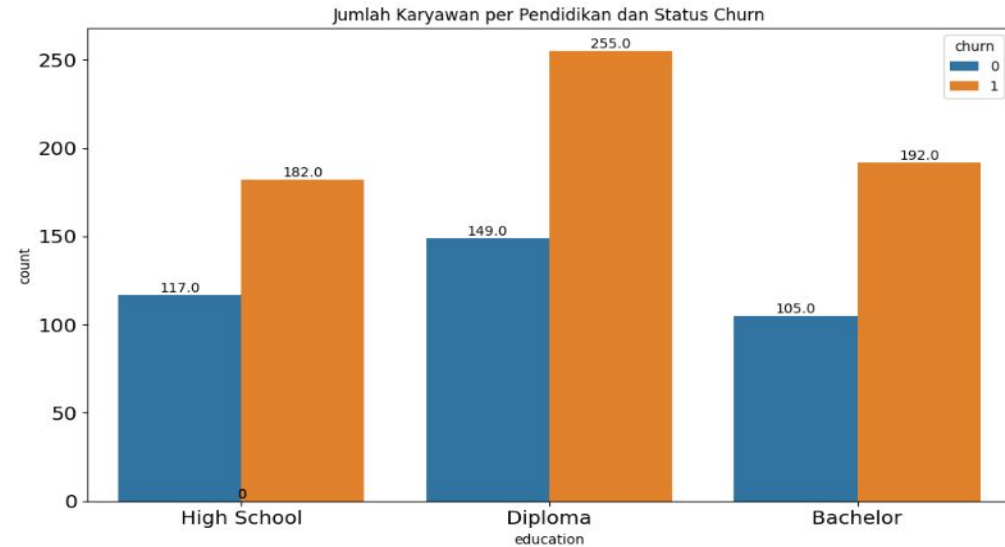


Dari line merah pada grafik dilihat garis melengkung , data baiknya menggunakan **non-linear model**.

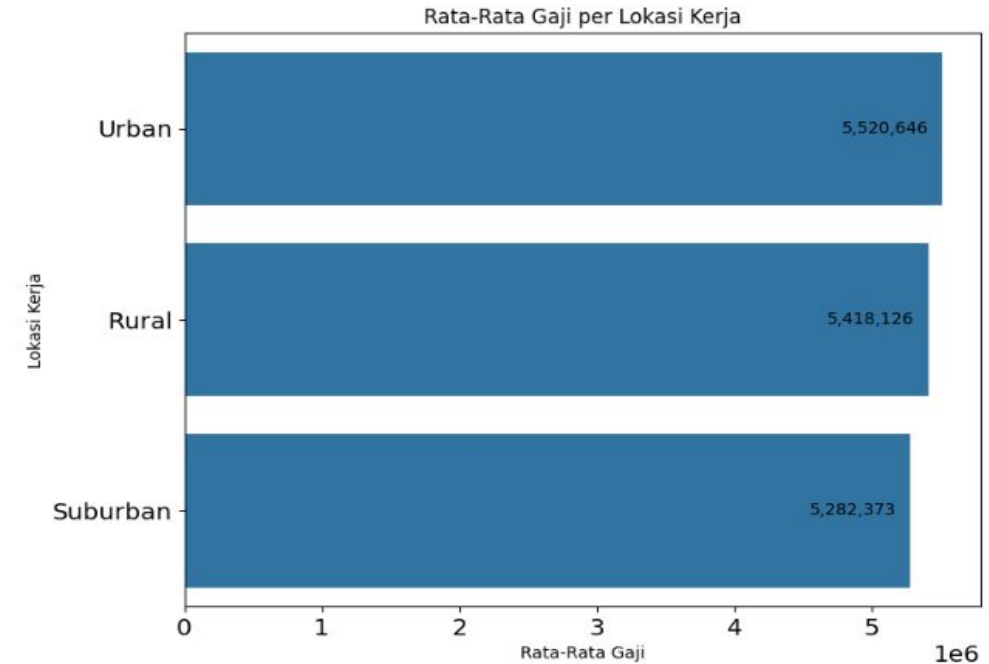
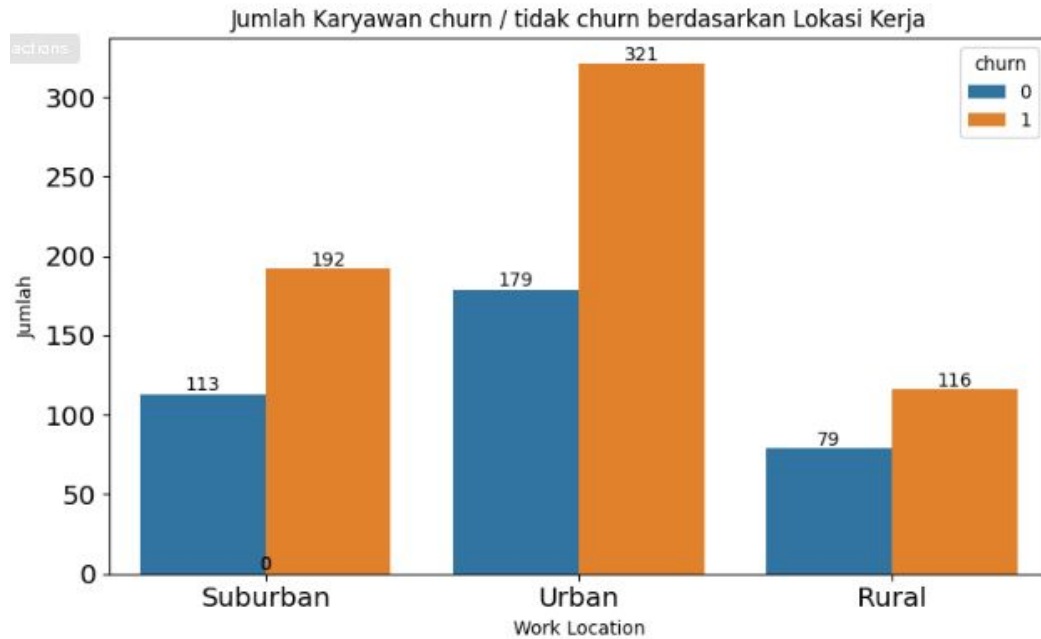
## Bivariate

Dari grafik di samping di ketahui, bahwa karyawan yang memiliki pendidikan di **diploma** cenderung **turnover**, ini karena dapat di lihat dari boxplot yang ada bahwa **target achievement** yang mereka capai **rendah** yang menyebabkan karyawan memilih resign.

**Solusi** yang dapat dilakukan : memberikan **pelatihan khusus** bila target achivment yang di dapati kurang dari 70%



## Bivariate



Dari grafik di atas dapat dilihat bahwa karyawan yang tinggalnya di **urban (perkotaan)** cenderung memilih **turnover**, ini disebabkan karna distribusi **salary** dari masing masing lokasi **tidak berbeda jauh**, sedangkan biaya hidup di setiap lokasi itu berbeda. Yang mana di perkotaan biaya hidup lebih tinggi

**Solusi** yang dapat dilakukan : mempertimbangkan **salary** sesuai dengan **tempat tinggal** dari karyawan



## Multivariate

churn	0.03	0.01	-0.02	-0.02	-0.28	0.18	-0.01	0.01	0.05	-0.21	-0.16	-0.10	1.00	0.18
distance_to_office_km	-0.01	-0.03	-0.02	-0.01	-0.02	-0.02	0.00	-0.00	0.00	0.03	0.01	-0.05	0.18	1.00
employee_id														
age														
experience_years														
monthly_target														
target_achievement														
working_hours_per_week														
overtime_hours_per_week														
salary														
commission_rate														
job_satisfaction														
manager_support_score														
company_tenure_years														
churn														
distance_to_office_km														

Kolerasi dari fitur – fitur yang ada **tidak saling berkolerasi secara kuat**, terlihat dari angkanya tidak ada yang > 0,7. Sehingga dapat menggunakan semua fitur tanpa ada yang perlu di hapus.



# DATA PREPROCESSING

---

TalentaHub

## Data Cleansing

Dari EDA yang telah dilakukan secara isi data dan tipe data serta persebarannya data **sudah bersih** dan dapat dilakukan tahap selanjutnya



# DATA PREPROCESSING

## Feature Engineering

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 24 columns):
#   Column                Non-Null Count  Dtype
---  -
0   employee_id           1000 non-null   int64
1   age                   1000 non-null   int64
2   gender                1000 non-null   object
3   education             1000 non-null   object
4   experience_years       1000 non-null   int64
5   monthly_target        1000 non-null   int64
6   target_achievement    1000 non-null   float64
7   working_hours_per_week 1000 non-null   int64
8   overtime_hours_per_week 1000 non-null   int64
9   salary                1000 non-null   int64
10  commission_rate       1000 non-null   float64
11  job_satisfaction      1000 non-null   int64
12  work_location         1000 non-null   object
13  manager_support_score  1000 non-null   int64
14  company_tenure_years  1000 non-null   float64
15  churn                 1000 non-null   int64
16  marital_status        1000 non-null   object
17  distance_to_office_km  1000 non-null   int64
18  churn_period          1000 non-null   object
19  target_gap            1000 non-null   float64
20  overwork_ratio        1000 non-null   float64
21  tenure_per_age        1000 non-null   float64
22  income_per_hour       1000 non-null   float64
23  experience_to_tenure   1000 non-null   float64
dtypes: float64(8), int64(11), object(5)
memory usage: 187.6+ KB
```

Menambah 5 variable pendukung

## Stage 2 : Model Development & Experimentation

# TUJUAN MODEL & METRIK EVALUASI

## Tujuan:

- \* Memprediksi apakah karyawan akan melakukan *churn* (resign) dalam waktu dekat dengan memberikan hasil segmentasinya.
- \* Memberi peringatan dini pada HR agar bisa intervensi lebih cepat.

## Metrik Evaluasi:

Accuracy  
Precision  
Recall  
F1-Score  
ROC-AUC

# SPLIT DATASET

Train-test split 80:20

```
✓ ▶ # pisahkan x,y dan split data test dan train
    # Pisahkan X dan y
    X = df_encoded.drop("churn", axis=1)
    y = df_encoded["churn"]

    # Split train-test
    X_train, X_test, y_train, y_test = train_test_split(
        X, y, test_size=0.2, stratify=y, random_state=42)
```

**X:** Fitur yang sudah di encoding

**y:** Label churn

# MODEL BASELINE

==== Logistic Regression ====

Confusion Matrix:

[[48 26]

[37 89]]

Classification Report:

	precision	recall	f1-score	support
0	0.56	0.65	0.60	74
1	0.77	0.71	0.74	126
accuracy			0.69	200
macro avg	0.67	0.68	0.67	200
weighted avg	0.70	0.69	0.69	200

Model untuk klasifikasi menggunakan model logistic regression sebagai baseline karna model sederhana

# EKSPERIMEN MODEL

	Model	Train Recall	Test Recall	Train Accuracy	Test Accuracy	Train ROC-AUC	Test ROC-AUC
0	Logistic Regression	1.00	1.00	0.63	0.63	0.50	0.50
1	Decision Tree	1.00	1.00	0.63	0.63	0.50	0.50
2	Random Forest	0.99	1.00	0.63	0.63	0.43	0.49
3	SVM	1.00	1.00	0.63	0.63	0.50	0.50
4	KNN	0.00	0.00	0.37	0.37	0.50	0.50

Dari semua pengujian yang dilakukan dari ke 4 model yang lebih baik adalah **Random Forest** karna **recallnya realistis** tetapi **ROC-AUC**nya masih rendah



# Evaluasi dan Interpretasi

Setelah dilakukan hyperparameter tuning maka hasilnya

Best ROC-AUC: 0.8812904236877719

Yang sebelumnya 0,49 mengalami kenaikan menjadi 0,88

**Random Forest** dipilih karna :

1. Kuat terhadap Overfitting
2. Menangani Fitur yang Banyak dan Kompleks
3. Performanya Umumnya Bagus

# Penyimpanan Model & Notebook

TalentaHub



## Stage 3 : Model Evaluation & Interpretability

# MODEL EVALUATION

Confusion Matrix:

```
[[56 18]
 [36 90]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.61	0.76	0.67	74
1	0.83	0.71	0.77	126
accuracy			0.73	200
macro avg	0.72	0.74	0.72	200
weighted avg	0.75	0.73	0.73	200

ROC-AUC Score: 0.8255040755040755

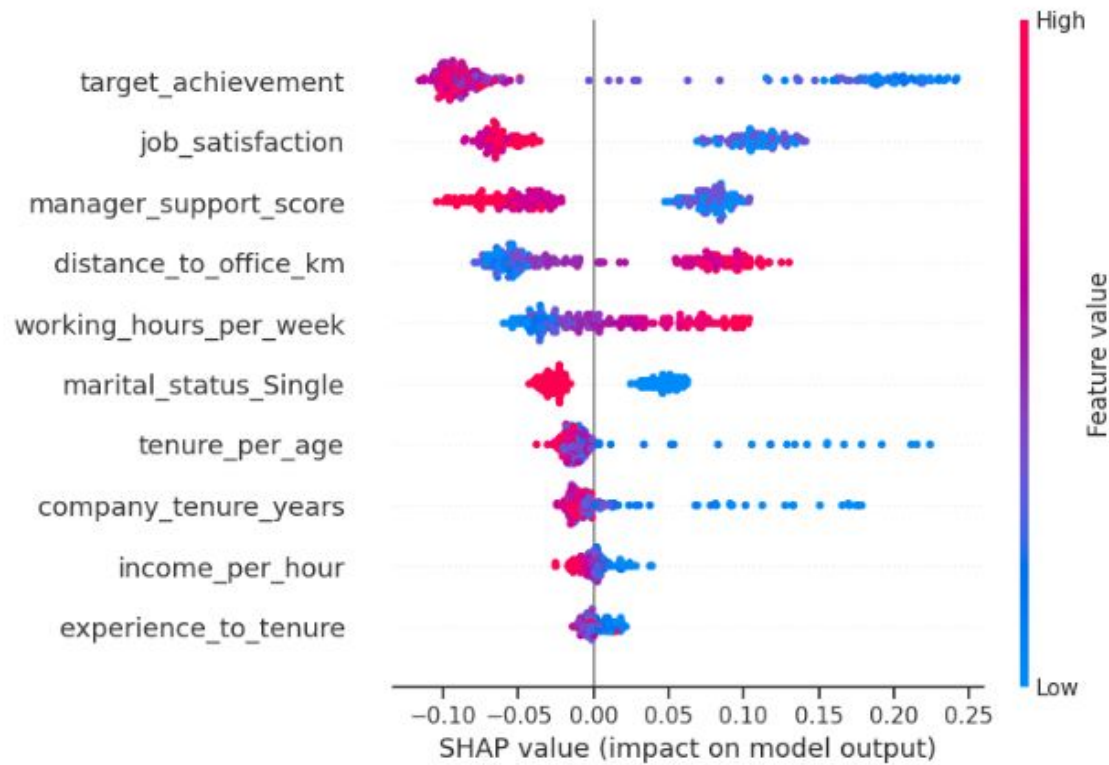
**True Negative (TN = 56):** Model benar memprediksi kelas 0 (tidak resign).

**False Positive (FP = 18):** Model salah prediksi 0 sebagai 1 (resign, padahal tidak).

**False Negative (FN = 36):** Model salah prediksi 1 sebagai 0 (tidak resign, padahal resign).

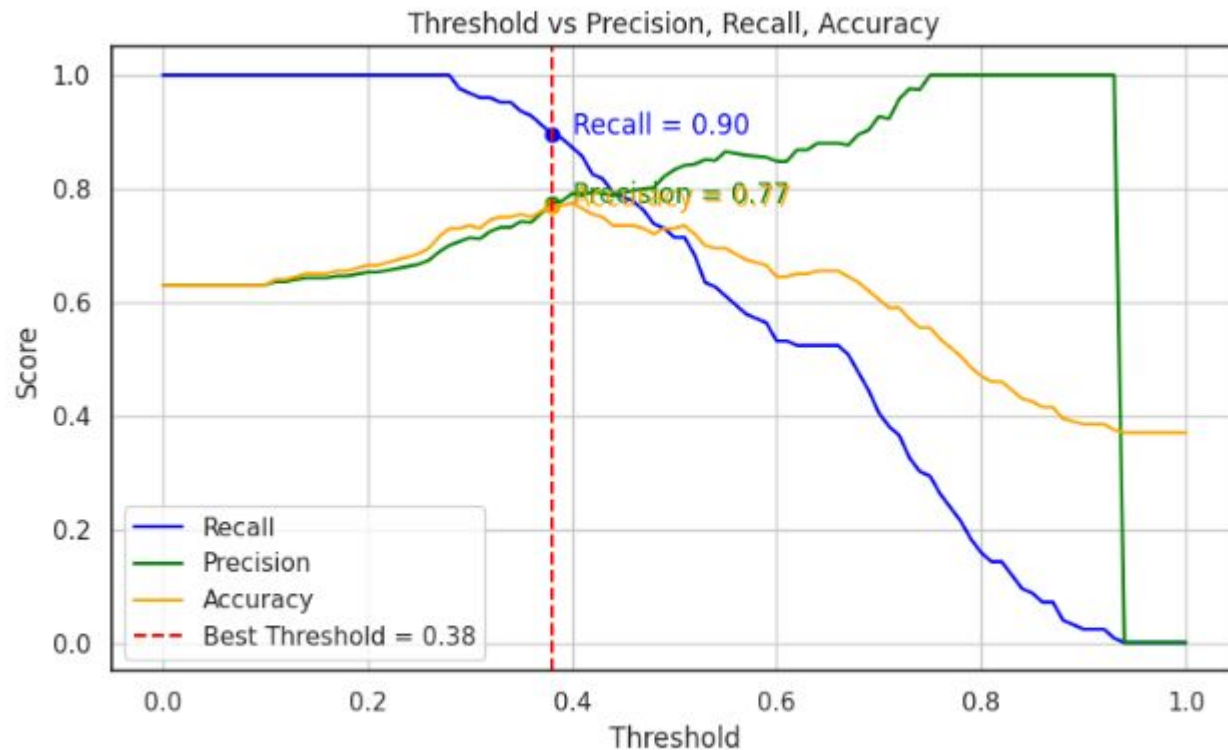
**True Positive (TP = 90):** Model benar memprediksi kelas 1 (resign).

# MODEL EXPLAINABILITY-SHAP



Setelah di lakukan shap terdapat 10 fitur terbaik yang dipakai untuk prediction deployment

# MODEL EXPLAINABILITY



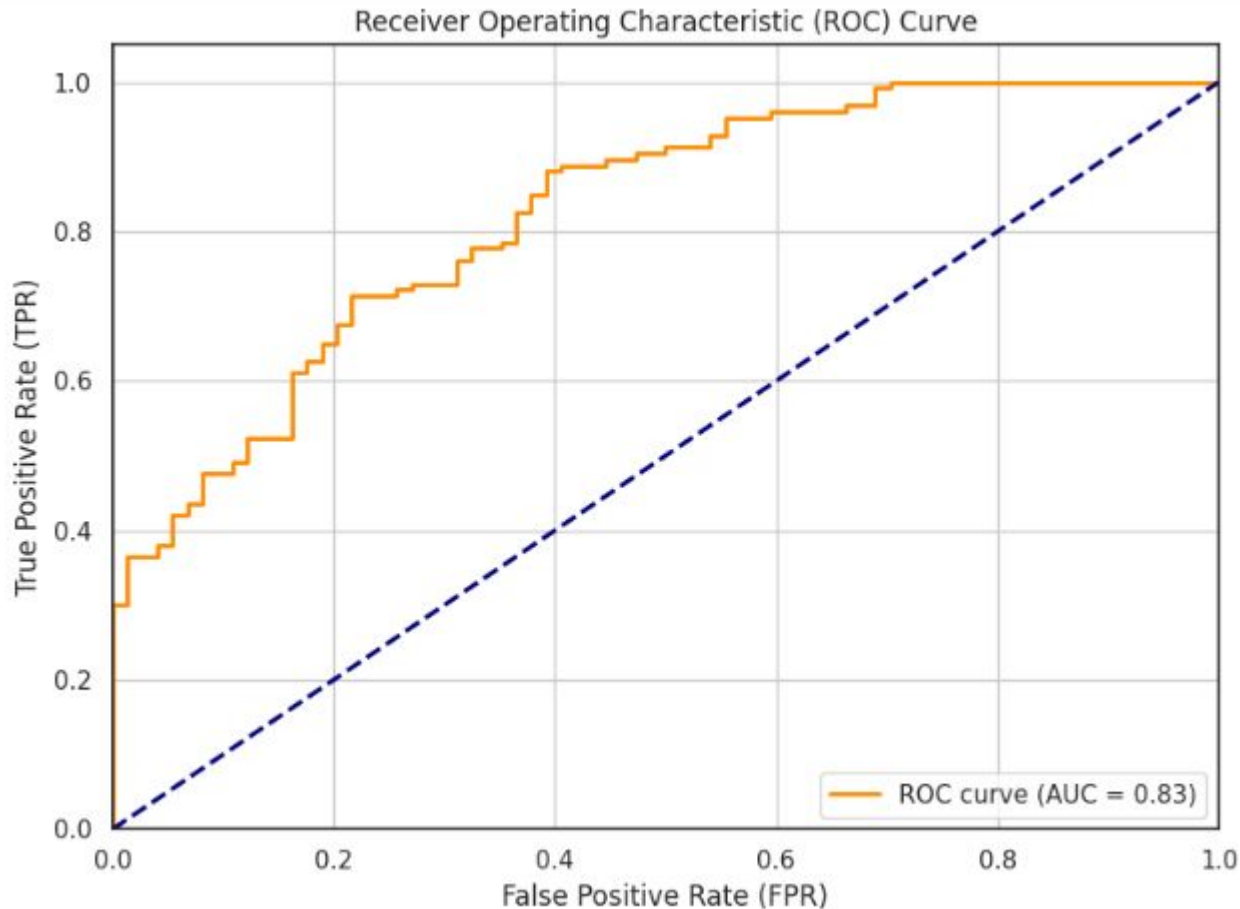
📌 Best Threshold (by F1-score): 0.38

Segmentasi turnover di bagi menjadi 3 kategori :

1. High >69%
2. Medium 38% - 69%
3. Low <38%

Dari Threshold yang di dapat , diatas 38% merupakan potensi karyawan turnover sehingga dari tim kami membuat segmentasi potensinya menjadi 3 kategori di dapat angka 69% dari hasil nilai tengah antara 38%-100%

# MODEL EXPLAINABILITY



Berdasarkan kurva ROC yang dihasilkan, model mencapai nilai AUC sebesar **0.83**, yang menunjukkan bahwa model memiliki **kemampuan yang baik dalam membedakan antara karyawan yang akan churn dan yang tidak**. Semakin besar nilai AUC, semakin baik model dalam klasifikasi, dan nilai ini berada pada kategori **baik**.

# BIAS & FAIRNESS

## 1. Model cenderung lebih baik memprediksi kelas 1 (resign):

**Precision** dan **F1-score** untuk kelas 1(churn) lebih tinggi.

Model lebih **percaya diri saat memprediksi kelas resign**, tapi tidak seakurat itu di kelas tidak resign.

## 2. Recall kelas 0 tinggi (0.76), tapi precision rendah (0.61)

Artinya: dari semua yang diprediksi **tidak resign**, cukup banyak yang salah (false positive).

Ini menunjukkan **kemungkinan bias terhadap prediksi positif (resign)**, karena model **menandai lebih banyak orang sebagai resign daripada sebenarnya**.

## 3. ROC-AUC Score: 0.83

Ini cukup baik secara umum, tapi **tidak menjamin tidak ada bias antar kelas**.

ROC-AUC bersifat **global**, tidak selalu mencerminkan **fairness antar kelas**.



# BUSINESS IMPACT

Perhitungan penurunan turnover dari model yang dibuat :

Turnover rate = 629 karyawan

Recall = 0,71

accuracy = 0,73

turnover karyawan awal = 629 orang

data total = 1000 orang

model menemukan karyawan turnover =

TP = recall \* turnover

= 0.71\*629

=447

**artinya 447 orang berhasil ditemukan**

intervensi = =447 \* accuracy

=447 \* 0.73

=327

**artinya 327 yang dapat dicegah resign**

turnover karyawan baru = 629 - 327 = 302 orang

turnover rate baru =

= resign yang baru/data \* 100%

= 302/1000 \* 100%

= 30.2% -

**sudah sesuai dengan kondisi ideal**

turnover rate awal

= turnover/data \* 100%




=629/1000 \* 100%

= 62.9 %



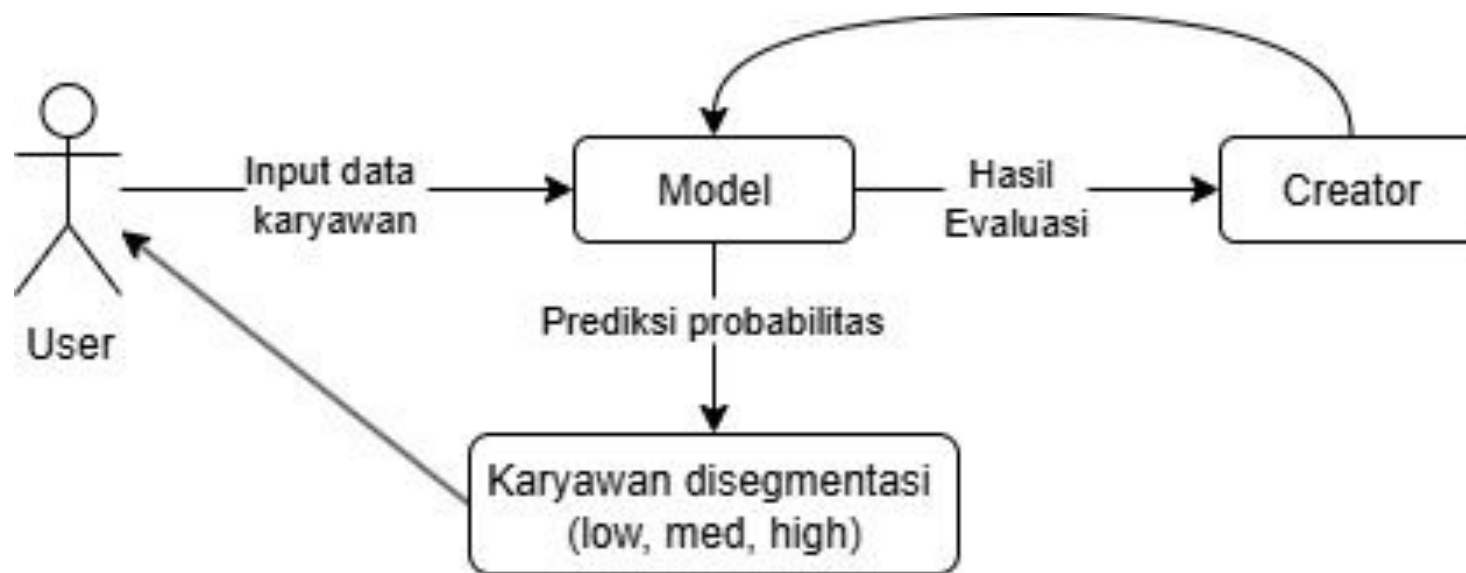
Penurunannya	=turnover rate lama - turnover rate baru = 62.9% - 30.2% = 32.9%
--------------	---

# BUSINESS IMPACT

Risk Level	% Kemungkinan Resign	Komponen Biaya Dampak	Estimasi Biaya/Orang	Penjelasan
 <b>High</b>	>69%	-Biaya rekrutmen baru -- Biaya pelatihan onboarding -- Hilangnya produktivitas -- Kehilangan pengetahuan organisasi (IP loss) -- Disengagement tim	60%	Berdasarkan Pendekatan Risk-Based Prioritization bahwa mengalokasikan sumber daya dan perhatian pada area yang memiliki risiko tertinggi terlebih dahulu
 <b>Medium</b>	38–69%	-Biaya rekrutmen (kemungkinan resign sedang) -- Pelatihan ulang sebagian -- Supervisi tambahan	30%	Belum tentu resign, tapi potensi kerugian tetap ada jika tidak dimitigasi
 <b>Low</b>	<38%	- Biaya monitoring & engagement- Survey kepuasan, wellness, dsb.	10%	Biaya retensi pasif, seperti wellbeing support, check-in, dan engagement plan

## Stage 4 : Deployment Evaluation & Interpretability

# DEPLOYMENT



Web aplikasi yang digunakan untuk memberikan segmentasi dan persentase karyawan turnover atau tidak menggunakan **streamlit**

# DEVELOPMENT


TalentaHub

×

Navigation

Prediction

Previous Predictions

 **Churn Prediction**

🎯 Target Achievement (%)

100

-

+

🕒 Working Hours per Week

35

-

+

😊 Job Satisfaction (1-5)

4.66

0.00

5.00

👤 Manager Support Score (1-5)

3.89

0.00

5.00

🏢 Company Tenures (years)

3.00

-

+

🚗 Distance to Office (km)

0

-

+

Tenure Age (years)

2.00

-

+

💰 Income (Hourly)

25000,00

-

+

🏆 Experience to Tenures (years)

1.00

-

+

👤 Marital Status


☒ Single

☐ Married

Predict


Churn Probability: 30%

Risk Group: Low

 Suggestion: Keep up the good work!

Made with Streamlit

[LINK DEVELOPMENT](#)

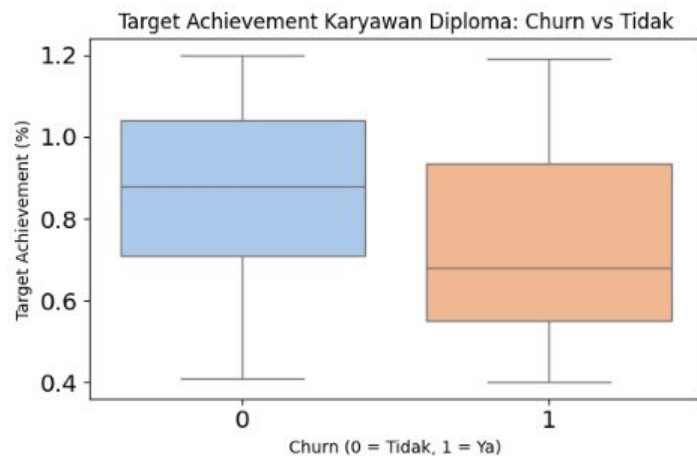
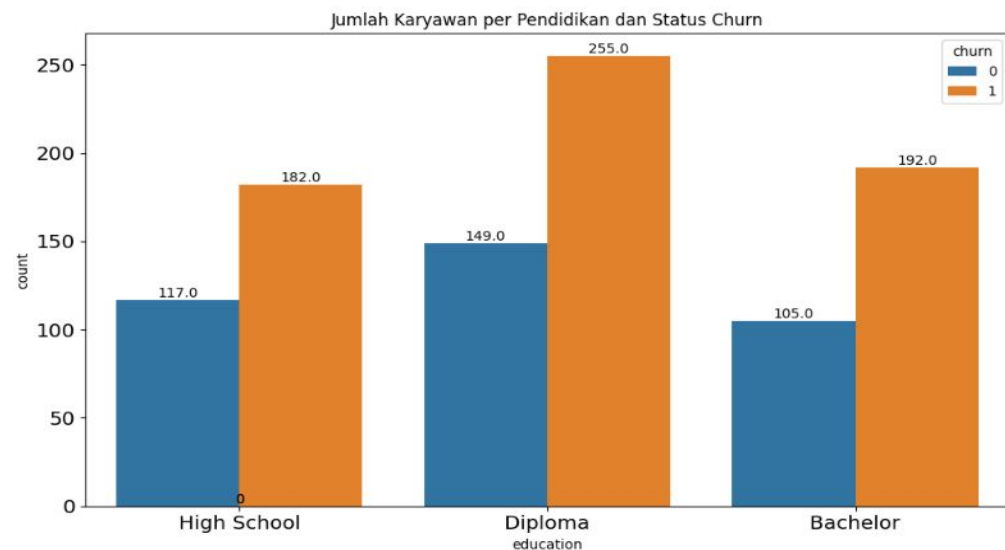
 **Rakamin**

# STRATEGI MONITORING MODEL

TalentaHub

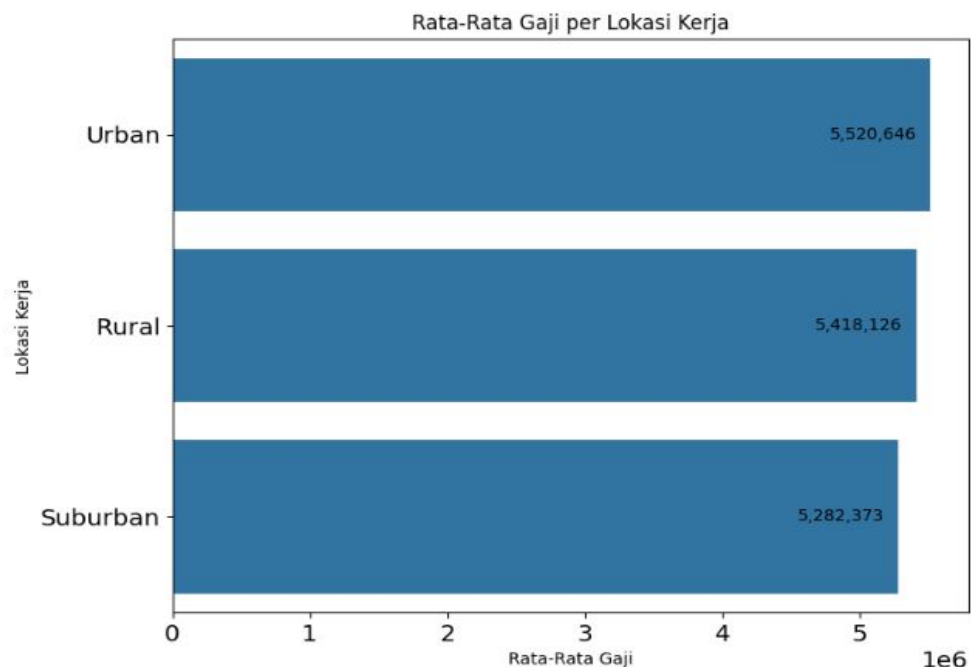
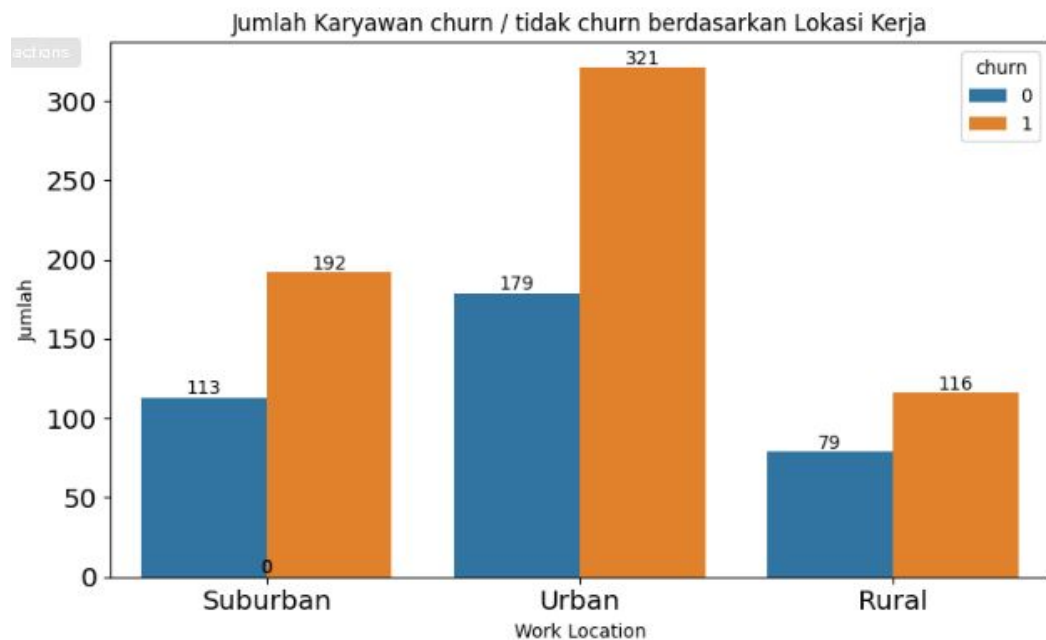
1. Setelah dilakukan deployment, simpan prediksi dan data aktual setiap minggu
2. Hitung ulang recall dan precision secara berkala (mingguan)
3. Plot distribusi input baru vs data training
4. Jika recall turun 15% lakukan retrain model
5. Minta feedback dari HR terkait intervensi model

# RECOMMENDATION



Pemberian **pelatihan** jika karyawan tidak dapat mencapai target

# RECOMMENDATION



**Pengalokasian gaji** sesuai umur  
pada setiap daerah tempat tinggal  
dapat di sesuaikan kembali



# RECOMMENDATION

TalentaHub

Tambahan fitur pendukung

Alasan resign

Jumlah Promosi

Jabatan

Status kepegawaian  
(kontrak/tetap)

Tanggal Masuk & Tanggal  
Resign

**THANK YOU**