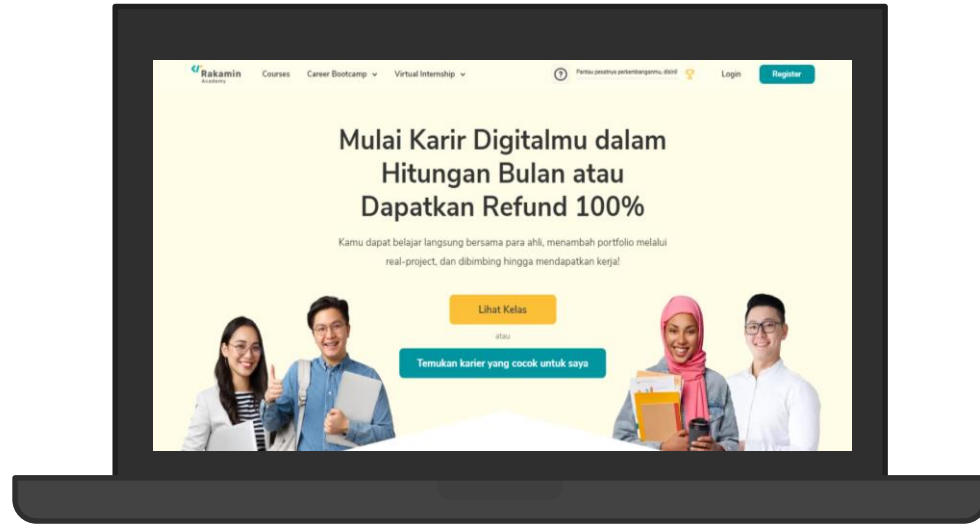




Improving Employee Retention by Predicting Employee Attrition Using Machine Learning

By Arya Octavian

...



Credits to Rakamin Academy

...

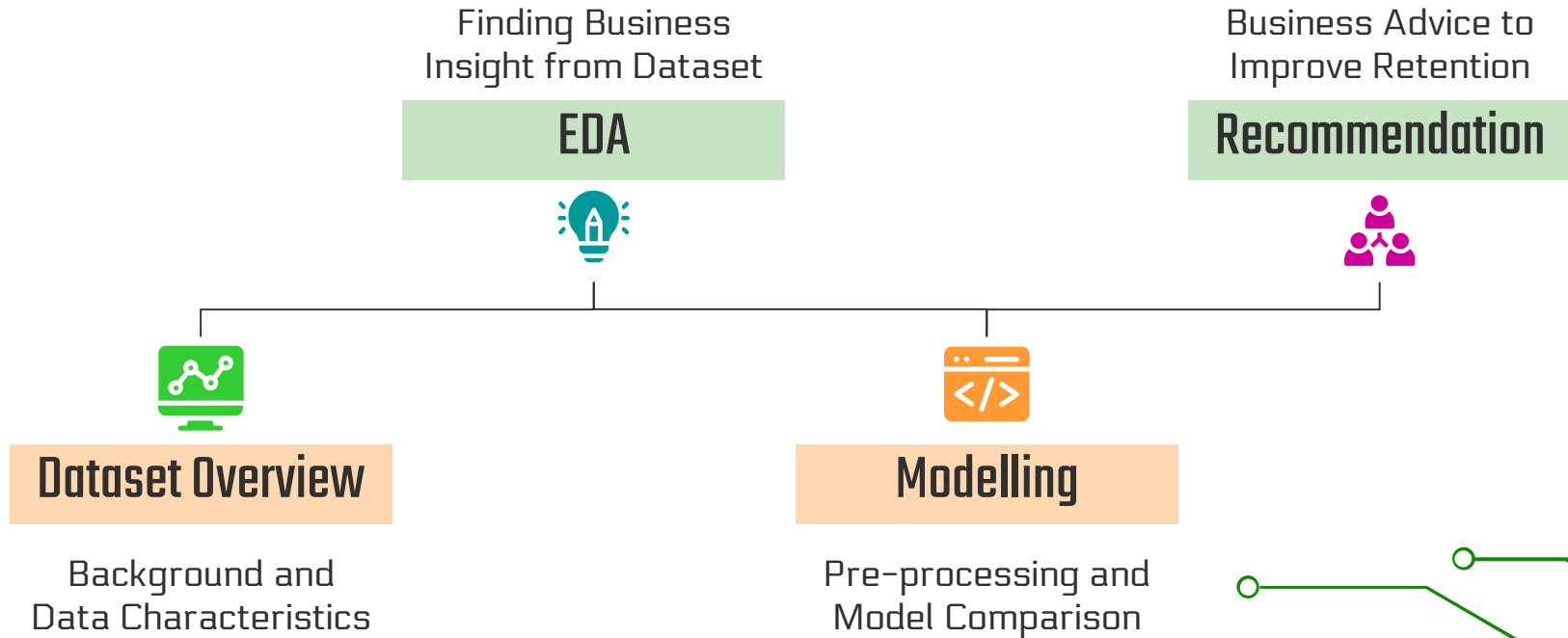
About Me

Greetings, you can call me Arya. I am 25 years old and I am a big fan of data analysis. Right now, I am taking a data science boot camp at Rakamin Academy. In the past, I used to work as a registration officer and data entry freelancer a couple of times. My long-term career goal is to contribute to the world as a business intelligence with great data analysis skills.

...



Timeline

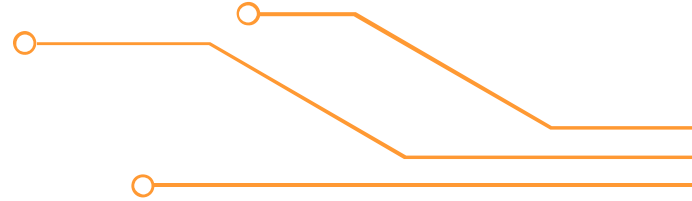


Dataset Overview

...



Background



01 Problem Statment

31% karyawan melakukan resign dari perusahaan.

02 Goal

Menurunkan potensi karyawan resign sebanyak 10%.

03 Objective

Klasifikasi karyawan yang berpotensi resign.

04 Business Metrics

Turn over rate atau persentase karyawan yang keluar.

...



Dataset

	Username	EnterpriseID	StatusPernikahan	JenisKelamin	StatusKepegawaian	Pekerjaan	JenjangKarir	PerformancePegawai
0	spiritedPorpoise3	111065	Belum_menikah	Pria	Outsource	Software Engineer (Back End)	Freshgraduate_program	Sangat_bagu
1	jealousGelding2	106080	Belum_menikah	Pria	FullTime	Data Analyst	Freshgraduate_program	Sangat_kuran
2	pluckyMuesli3	106452	Menikah	Pria	FullTime	Software Engineer (Front End)	Freshgraduate_program	Bagu
3	stressedTruffle1	106325	Belum_menikah	Pria	Outsource	Software Engineer (Front End)	Freshgraduate_program	Bagu
4	shyTermite7	111171	Belum_menikah	Wanita	FullTime	Product Manager	Freshgraduate_program	Bagu
...
282	dopeySheep0	106034	Belum_menikah	Wanita	FullTime	Data Engineer	Mid_level	Bagu
283	yearningPorpoise4	106254	-	Wanita	FullTime	Product Design (UI & UX)	Freshgraduate_program	Bias

Dataset memiliki 287 baris dan 25 kolom

5 kolom numerik

SkorSurveyEngagement
SkorKepuasanPegawai
JumlahKeikutsertaanProjek
JumlahKeterlambatanSebulanTerakhir
JumlahKetidakhadiran

20 kolom kategorik

JenjangKarir, StatusKepegawaian, dll

...

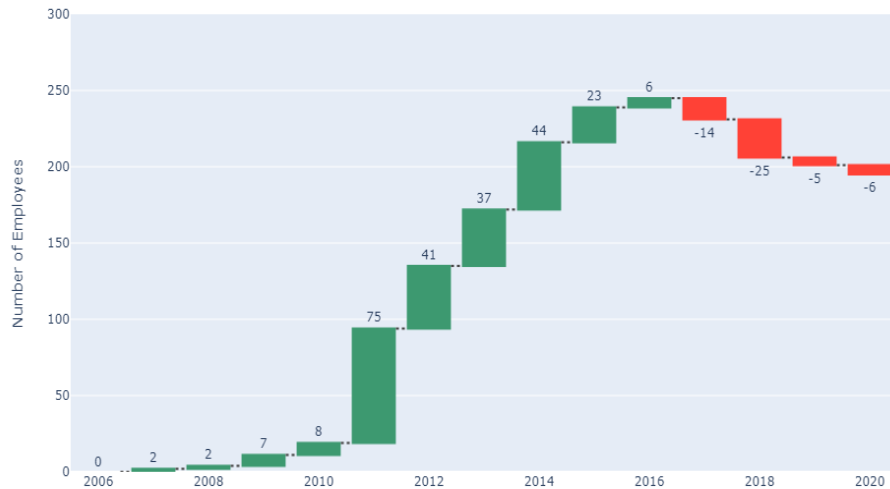
EDA

(Exploratory Data Analysis)

...

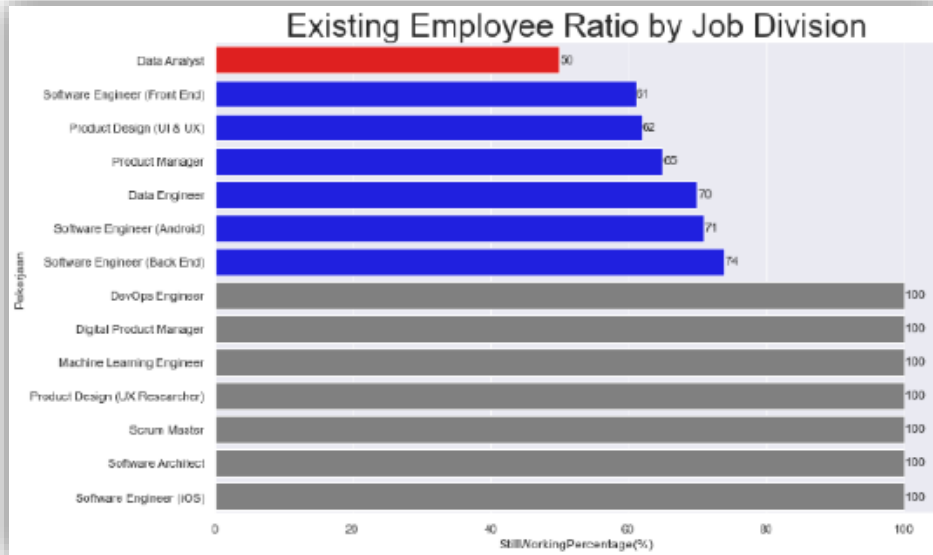


Number of Employee Changes Per Year



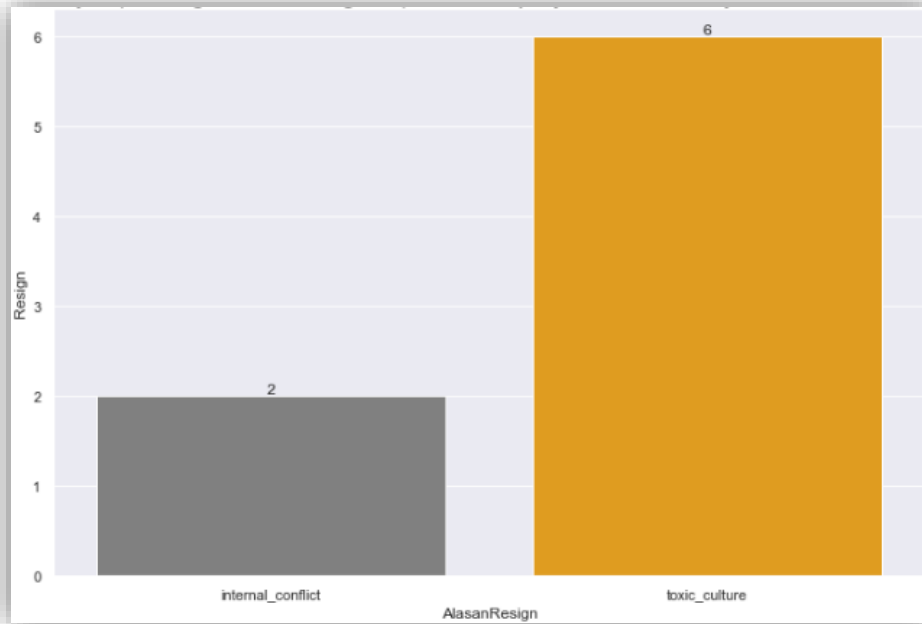
- Mulai tahun 2017 jumlah SDM mengalami pengurangan yang lebih banyak daripada yang direkrut hal ini disebabkan karena permasalahan internal.
- Secara keseluruhan turn over rate perusahaan dari awal beroperasi hingga 2021 adalah 45%.
- Perusahaan memahami divisi mana yang paling krusial apabila belum diisi kursi kosong jabatannya.

Existing Employee Ratio by Job Division



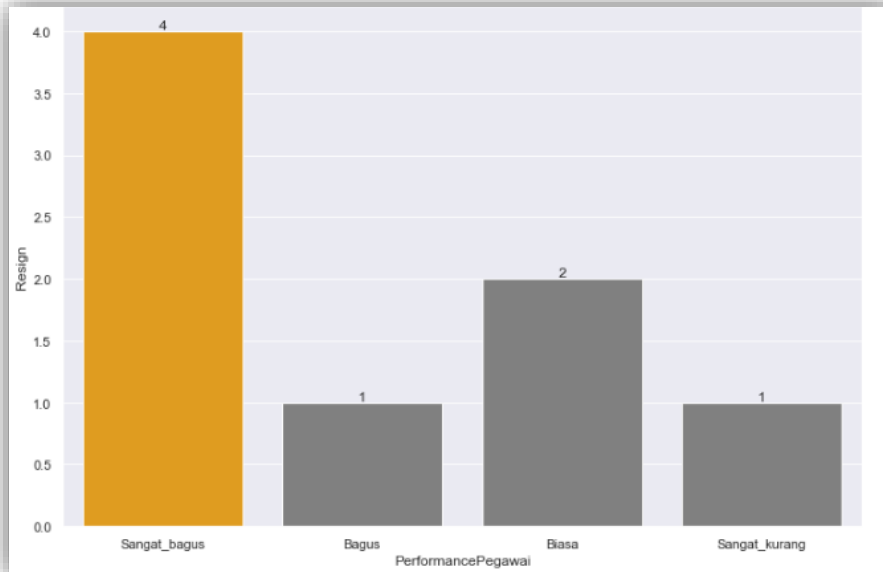
- Karena hampir semua alasan resign pada dataset adalah terkait toxic culture, perusahaan dapat mempertahankan karyawan dengan bar warna biru dengan **rutin mengadakan acara akhir bulan seperti program aktivitas yang menyenangkan diikuti oleh seluruh karyawan.**
- Perusahaan harus mengoptimalkan retensi karyawan yang bekerja pada divisi dengan rasio pekerja 100%. Hal ini disebabkan oleh **jumlah karyawannya sedikit yang berarti akan sulit untuk mendapatkan SDM serupa.**
- **Divisi data analyst memiliki rasio resign paling besar.** Dikarenakan hal ini menjadi krusial pada performa divisi, maka perlu dicari tahu karakteristik seperti apa yang menyebabkan karyawan data analyst mengundurkan diri.

Data Analyst Total Employees Attrition by Reason for Resignation



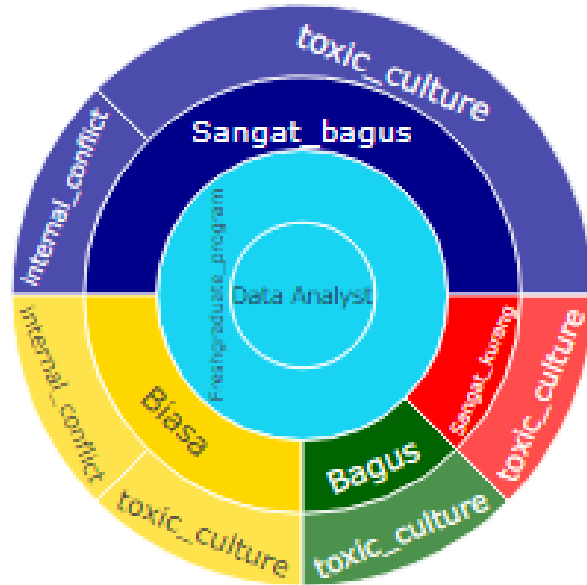
- Semua karyawan data analyst yang resign disebabkan oleh toxic culture dan internal conflict. Kedua hal ini merupakan permasalahan yang serupa.
- Toxic culture bisa datang dari berbagai hal seperti deadline yang tidak wajar maupun adanya intimidasi antar karyawan.
- Perusahaan perlu menerapkan workflow yang jelas agar tidak ada celah bagi atasan divisi untuk memberikan tugas secara tidak wajar. Penerapan scrum dapat mengurangi konflik pada karyawan karena pekerjaan. Hal ini karena scrum dalam menetapkan KPI perlu atas keputusan bersama. Selain itu, scrum master juga dapat membantu mengevaluasi hubungan antar karyawan.

Data Analyst Total Employees Attrition by Performance



- Data Analyst yang resign didominasi oleh karyawan dengan performa sangat bagus. Hal ini menjadi loss yang besar bagi SDM perusahaan.
- Selain mengevaluasi toxic culture dalam divisi, perusahaan juga dapat menerapkan sistem rolling tugas sehingga karyawan yang performanya kurang dapat menemukan kelebihanannya di tugas yang lain.
- Dengan adanya budaya kerja yang selalu mengapresiasi setiap effort yang karyawan lakukan pasti karyawan akan merasa termotivasi untuk bekerja lebih produktif dan dapat meredam konflik yang muncul dalam bekerja.

Sunburst Chart



- Secara kesimpulan semua toxic culture penyebab resign terjadi pada seluruh performa karyawan di divisi data analyst.
- Perusahaan hari mencari tahu apa yang menjadi sumber toxic culture yang dialami oleh seluruh karyawan agar tidak lebih jauh menurunkan organization health index. Tidak transparan, intimidasi, dan mikro manajemen merupakan faktor yang disarankan untuk dievaluasi.
- Salah satu cara paling efektif dalam mencegah toxic culture adalah dengan mencari informasi umpan balik dari karyawan. Hal ini juga dapat dimanfaatkan sebagai fitur baru di kemudian hari untuk meningkatkan kualitas model machine learning.

...

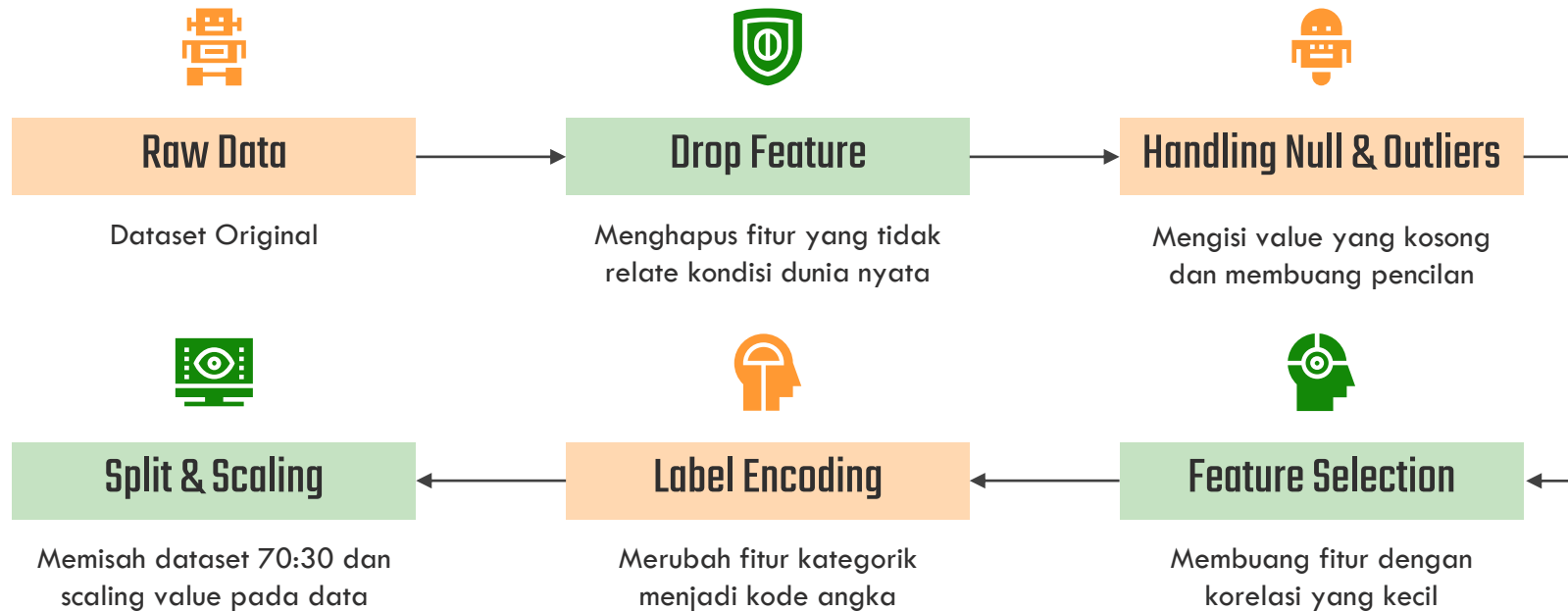
Modelling

...





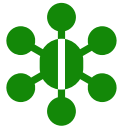
Preprocessing



...



Modelling Workflow



Oversampling & Modelling

Oversampling dengan SMOTE dan menggunakan 5 algoritma berbeda untuk proses machine learning



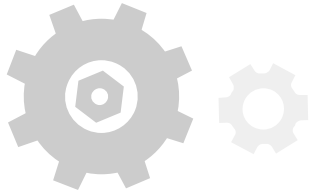
Tuning Hyperparameter

Melakukan hyperparameter tuning pada semua algoritma yang dipakai



Model Comparison

Merangkum metrik evaluasi dari semua algoritma dan memilih yang terbaik.



...



Model Comparison After Tuning

Model dengan algoritma XGBoost dipilih karena:

- Memiliki AUC-Proba Train yang tidak underfitting dan tidak terlalu overfitting.
- Nilai AUC-Proba Test yang tinggi dengan jarak 0.73% dari data Trainnya.
- Dikarenakan fokusnya adalah tidak ingin ada opportunity loss dalam memprediksi karyawan yang resign maka nilai Recall juga menjadi acuan.

Model Comparison

Tuned Logistic Regression	100.00%	99.74%	84.62%	100.00%	91.67%	95.29%
Tuned Decision Tree	97.44%	54.76%	26.92%	35.00%	30.43%	62.35%
Tuned Random Forest	86.99%	79.86%	53.85%	66.67%	59.58%	77.65%
Tuned Adaboost	100.00%	99.48%	84.62%	100.00%	91.67%	95.29%
Tuned XGBoost	99.82%	99.09%	96.15%	96.15%	96.15%	97.65%
	AUC-Proba Train	AUC-Proba Test	Recall	Precision	F1	Accuracy

...

For full code:

[Link Jupyter Notebook](#)

...

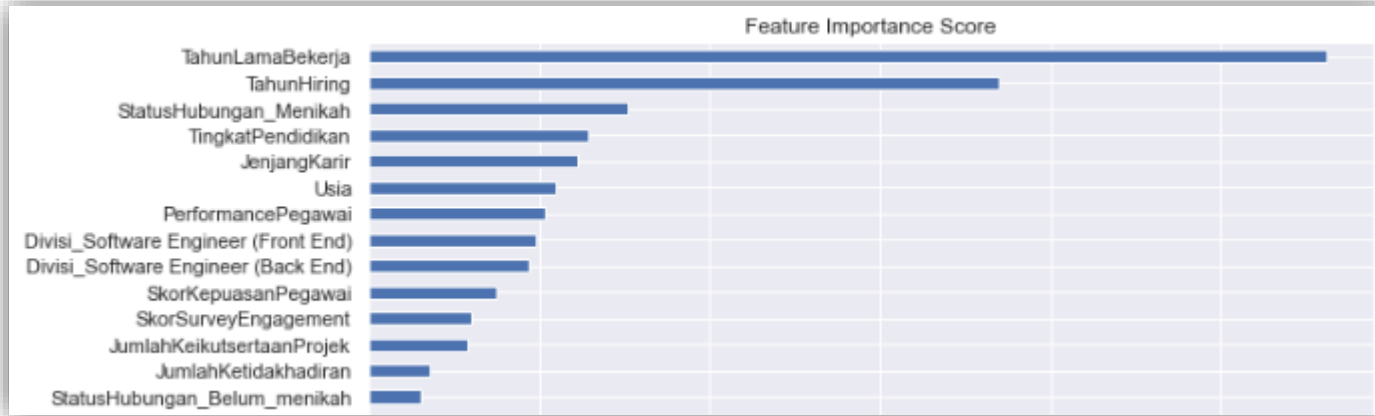
You can visit my projects:

[Link Github Profile](#)

Recommendation



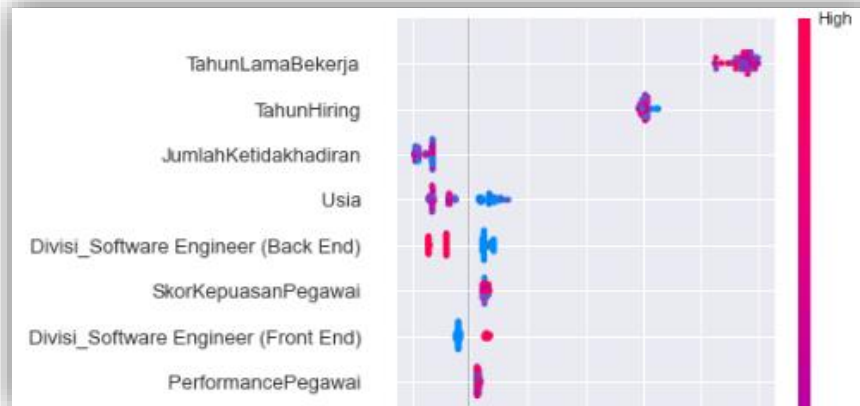
Feature Importance



- Tahun lama bekerja dan tahun hiring merupakan fitur yang paling berpengaruh terhadap kemungkinan resign karyawan.
- Hal ini menggambarkan bahwa **seberapa lamanya karyawan bekerja untuk perusahaan** akan mempengaruhi keputusan karyawan untuk resign atau tidak dengan motivasi dari berbagai alasan.
- Dengan begitu perusahaan dapat memperhatikan kelompok-kelompok karyawan berdasarkan sudah berapa lama (dalam tahun) mereka bekerja untuk perusahaan.

...

Shap Values



- Persebaran data yang dihasilkan model terlihat sangat acak dalam menentukan karyawan akan resign atau tidak.
- Namun dari shap value dapat diperoleh bahwa **semakin muda usia karyawan** maka semakin berpotensi untuk resign dari perusahaan.

...

Business Recommendation



- Perusahaan memfokuskan retensi karyawan pada karyawan yang umurnya lebih muda.
- Rutin memberikan apresiasi loyalitas seperti bonus gaji untuk karyawan yang > 9 tahun bekerja untuk perusahaan.
- Menerapkan workflow yang konsisten kepada semua divisi sambil mengevaluasi tingkat toxic culture di perusahaan.
- Melakukan rekrutmen minimal 5% dari total karyawan untuk setiap divisi yang terdapat resign.
- Melakukan evaluasi kepada karyawan yang telah tidak masuk > 10 kali atau karyawan yang terlambat > 0 kali tiap bulan.
- Memperbanyak kegiatan proyek agar minimal setiap karyawan pernah mengikuti minimal 1 proyek.

Research Suggestions



Add More Features

Menambahkan fitur seperti:

Feedback karyawan

Detail jenjang karir divisi

Gaji bulanan

Jumlah cuti yang dipakai

...



Koreksi Dataset

Mengevaluasi konten dataset yang salah akibat human error.

Seperti tanggal pada fitur resign lebih dahulu terjadi daripada tanggal fitur penilaian

...





THANKS!

DO YOU HAVE ANY QUESTIONS?

arya.octavian23@gmail.com

<https://github.com/aryaoctav>

<https://www.linkedin.com/in/arya-octavian/>

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik** and illustrations by **Storyset**