

Improving Employee Retention by Predicting Employee Attrition Using Machine Learning

Supported by:
Rakamin Academy
Career Acceleration School
www.rakamin.com



Created by:

Rheza Paleva Uyanto

uyantorheza@gmail.com

<https://www.linkedin.com/in/rheza-uyanto/>

Lulus dari Program Pendidikan Profesi Apoteker Universitas Surabaya pada tahun 2017, memiliki pengalaman praktik kefarmasian di Rumah Sakit selama 4 tahun. Kemampuan komunikasi dan managerial yang baik, dan mampu bekerja sama dalam tim ataupun secara mandiri. Sangat termotivasi dalam bidang mentoring dan pengembangan diri. Mampu dalam penggunaan Microsoft Office. Memiliki ketertarikan dalam bidang Data Analitik, sehingga mengikuti Rakamin Bootcamp Data Science Batch 24. Melalui portofolio ini, saya akan memprediksi karyawan yang resign untuk meningkatkan retensi karyawan di sebuah perusahaan dengan menggunakan Machine Learning.

Overview

“Sumber daya manusia (SDM) adalah aset utama yang perlu dikelola dengan baik oleh perusahaan agar tujuan bisnis dapat tercapai dengan efektif dan efisien. Pada kesempatan kali ini, kita akan menghadapi sebuah permasalahan tentang sumber daya manusia yang ada di perusahaan. Fokus kita adalah untuk mengetahui bagaimana cara menjaga karyawan agar tetap bertahan di perusahaan yang ada saat ini yang dapat mengakibatkan bengkaknya biaya untuk rekrutmen karyawan serta pelatihan untuk mereka yang baru masuk. Dengan mengetahui faktor utama yang menyebabkan karyawan tidak merasa, perusahaan dapat segera menanggulanginya dengan membuat program-program yang relevan dengan permasalahan karyawan.”

Data Preprocessing

- Dataset : Improving Employee Retention by Predicting Employee Attrition Using Machine Learning.xlsx - hr_data.csv
- Terdiri dari 25 kolom dan 287 baris

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 287 entries, 0 to 286
Data columns (total 25 columns):
 #   Column           Non-Null Count Dtype  
 --- 
 0   Username        287 non-null    object  
 1   EnterpriseID   287 non-null    int64   
 2   StatusPernikahan 287 non-null    object  
 3   JenisKelamin    287 non-null    object  
 4   StatusKepegawaian 287 non-null    object  
 5   Pekerjaan       287 non-null    object  
 6   JenjangKarir    287 non-null    object  
 7   PerformancePegawai 287 non-null    object  
 8   AsalDaerah      287 non-null    object  
 9   HiringPlatform   287 non-null    object  
 10  SkorSurveyEngagement 287 non-null    int64   
 11  SkorKepuasanPegawai 282 non-null    float64 
 12  JumlahKeikutsertaanProjek 284 non-null    float64 
 13  JumlahKeterlambatanSebulanTerakhir 286 non-null    float64 
 14  JumlahKetidakhadiran 281 non-null    float64 
 15  NomorHP          287 non-null    object  
 16  Email             287 non-null    object  
 17  TingkatPendidikan 287 non-null    object  
 18  PernahBekerja    287 non-null    object  
 19  IkutProgramLOP   29 non-null     float64  
 20  AlasanResign    221 non-null    object  
 21  TanggalLahir     287 non-null    object  
 22  TanggalHiring    287 non-null    object  
 23  TanggalPenilaianKaryawan 287 non-null    object  
 24  TanggalResign   287 non-null    object  
dtypes: float64(5), int64(2), object(18)
memory usage: 56.2+ KB
```

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

Data Preprocessing

- Terdapat beberapa null value dan cara handlingnya:
 - SkorKepuasanPegawai** : 5 NaN diisi median
 - JumlahKeikutsertaanProjek** : 3 NaN diisi median
 - JumlahKeterlambatanSebulanTerakhir** : 1 NaN, diisi median
 - JumlahKetidakhadiran** : 6 NaN, diisi median
 - AlasanResign** : 66 NaN, diisi 'masih bekerja'
- Fitur yang berbeda:
 - StatusPernikahan**, tidak ada null value : 3 terisi '–', diganti ke 'Lainnya', karena bermakna sama.
- Fitur yang di drop:
 - PernahBekerja**, tidak ada null value tetapi ada 1 baris yang berisi 'yes' sedangkan lainnya 1, 1 dan yes bermakna sama sehingga hanya terdiri dari 1 nilai, kolom didrop
 - IkutProgramLOP** terdapat 258 null dari 287 value, lebih dari 50%, sehingga kolom didrop.
 - Username**, karena kardinalitas tinggi
 - NomorHP**, karena kardinalitas tinggi
 - Email**, karena kardinalitas tinggi

```
df.isna().sum()

Username 0
EnterpriseID 0
StatusPernikahan 0
JenisKelamin 0
StatusKepegawaian 0
Pekerjaan 0
JenjangKarir 0
PerformancePegawai 0
AsalDaerah 0
HiringPlatform 0
SkorSurveyEngagement 0
SkorKepuasanPegawai 5
JumlahKeikutsertaanProjek 3
JumlahKeterlambatanSebulanTerakhir 1
JumlahKetidakhadiran 6
NomorHP 0
Email 0
TingkatPendidikan 0
PernahBekerja 0
IkutProgramLOP 258
AlasanResign 66
TanggalLahir 0
TanggalHiring 0
TanggalPenilaianKaryawan 0
TanggalResign 0
dtype: int64
```

Data Preprocessing

- Setelah dilakukan handling missing value, dan dan dropping feature, feature yang tersedia antara lain :
 - Terdiri dari 20 kolom, dan 287 baris
 - Tidak terdapat nilai missing value
 - Tidak terdapat nilai duplikat

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 287 entries, 0 to 286
Data columns (total 20 columns):
 #   Column          Non-Null Count  Dtype  
--- 
 0   EnterpriseID   287 non-null    object  
 1   StatusPernikahan 287 non-null    object  
 2   JenisKelamin    287 non-null    object  
 3   StatusKepegawaian 287 non-null    object  
 4   Pekerjaan        287 non-null    object  
 5   JenjangKarir    287 non-null    object  
 6   PerformancePegawai 287 non-null    object  
 7   AsalDaerah      287 non-null    object  
 8   HiringPlatform   287 non-null    object  
 9   SkorSurveyEngagement 287 non-null    int64  
 10  SkorKepuasanPegawai 287 non-null    float64 
 11  JumlahKeikutsertaanProjek 287 non-null    float64 
 12  JumlahKeterlambatanSebulanTerakhir 287 non-null    float64 
 13  JumlahKetidakhadiran 287 non-null    float64 
 14  TingkatPendidikan   287 non-null    object  
 15  AlasanResign     287 non-null    object  
 16  TanggalLahir     287 non-null    object  
 17  TanggalHiring    287 non-null    object  
 18  TanggalPenilaianKaryawan 287 non-null    object  
 19  TanggalResign    287 non-null    object  
dtypes: float64(4), int64(1), object(15)
memory usage: 45.0+ KB
```

Annual Report on Employee Number Changes

- Untuk mendapat grafik pertumbuhan kartawan setiap tahun, diperlukan proses agregasi data
 - Melakukan ekstrasi data tanggal_hiring untuk mendapatkan tahun_hiring
 - Melakukan aggregasi dengan groupby sehingga didapatkan jumlah karyawan yang dihireng pertahun.
 - Melakukan hal yang sama untuk mendapatkan tahun_resign
 - Melakukan outer join, untuk menggabungkan tabel hiring dan resign

tahun_hiring	jumlah_hiring
0	2006
1	2007
2	2008
3	2009
4	2010
5	2011
6	2012
7	2013
8	2014
9	2015
10	2016
11	2017
12	2018

tahun_resign	jumlah_resign
0	2013
1	2014
2	2015
3	2016
4	2017
5	2018
6	2019
7	2020

tahun_hiring	jumlah_hiring	tahun_resign	jumlah_resign
0	2006	1.0	NaN
1	2007	2.0	NaN
2	2008	2.0	NaN
3	2009	7.0	NaN
4	2010	8.0	NaN
5	2011	76.0	NaN
6	2012	41.0	NaN
7	2013	43.0	2013
8	2014	56.0	2014
9	2015	31.0	2015
10	2016	14.0	2016
11	2017	5.0	2017
12	2018	1.0	2018
13	NaN	NaN	2019
14	NaN	NaN	2020

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

Annual Report on Employee Number Changes

- Selanjutnya dilakukan :

- Mengisi kolom null value pada kolom tahun_hiring dan tahun_resign, sesuai dengan tahun nya.

	tahun_hiring	jumlah_hiring	tahun_resign	jumlah_resign
0	2006	1.0	NaN	NaN
1	2007	2.0	NaN	NaN
2	2008	2.0	NaN	NaN
3	2009	7.0	NaN	NaN
4	2010	8.0	NaN	NaN
5	2011	76.0	NaN	NaN
6	2012	41.0	NaN	NaN
7	2013	43.0	2013	5.0
8	2014	56.0	2014	12.0
9	2015	31.0	2015	8.0
10	2016	14.0	2016	8.0
11	2017	5.0	2017	19.0
12	2018	1.0	2018	26.0
13	NaN	NaN	2019	5.0
14	NaN	NaN	2020	6.0

- Menghitung total karyawan kumulatif (dengan menambahkan total karyawan sebelumnya ditambahkan dengan jumlah_hiring dan dikurangi jumlah_resign
- Menghitung perubahan jumlah karyawan pertahun

	tahun_hiring	jumlah_hiring	tahun_resign	jumlah_resign	total_karyawan	perubahan
0	2006	1.0	2006	0.0	1.0	1.0
1	2007	2.0	2007	0.0	3.0	2.0
2	2008	2.0	2008	0.0	5.0	2.0
3	2009	7.0	2009	0.0	12.0	7.0
4	2010	8.0	2010	0.0	20.0	8.0
5	2011	76.0	2011	0.0	96.0	76.0
6	2012	41.0	2012	0.0	137.0	41.0
7	2013	43.0	2013	5.0	175.0	38.0
8	2014	56.0	2014	12.0	219.0	44.0
9	2015	31.0	2015	8.0	242.0	23.0
10	2016	14.0	2016	8.0	248.0	6.0
11	2017	5.0	2017	19.0	234.0	-14.0
12	2018	1.0	2018	26.0	209.0	-25.0
13	2019	0.0	2019	5.0	204.0	-5.0
14	2020	0.0	2020	6.0	198.0	-6.0

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

Annual Report on Employee Number Changes

Dinamika Pertumbuhan Karyawan tahun 2006-2020

Jumlah Karyawan



Berdasarkan waterfall plot, pada 10 tahun pertama (2006-2016), jumlah karyawan yang direkrut terus bertambah dibandingkan dengan jumlah karyawan yang resign.

Pertumbuhan jumlah karyawan yang paling tinggi sebanyak **76 orang** terjadi pada tahun 2011. Sampai pada tahun 2016, pertumbuhan jumlah karyawan berhenti dengan bertambahnya 6 orang. Pada akhir tahun 2016, jumlah karyawan sebanyak 248 orang

Setelah tahun 2016, karyawan yang resign lebih banyak. Karyawan yang resign paling banyak terjadi pada tahun 2018, yakni berkurang **25 orang**.

Sampai tahun 2020, total karyawan yang resign terus bertambah, sehingga akhir tahun 2020, jumlah karyawan sebanyak 198 orang. Dari tahun 2016 sampai akhir tahun 2020, jumlah karyawan berkurang sebanyak 20,16% atau sebanyak **50 orang**.

Resign Reason Analysis for Employee Attrition Management Strategy

- Untuk mendapat persentase resign dari setiap bidang, diperlukan proses agregasi data
 - Melakukan aggregasi jumlah karyawan yang resign dengan groupby per jenis pekerjaan.
 - Melakukan aggregasi jumlah karyawan yang belum resign dengan groupby per jenis pekerjaan.
 - Melakukan outer join, untuk menggabungkan tabel hiring dan resign

Pekerjaan	jumlah_belum_resign
0 Data Analyst	8
1 Data Engineer	7
2 DevOps Engineer	3
3 Digital Product Manager	2
4 Machine Learning Engineer	2
5 Product Design (UI & UX)	15
6 Product Design (UX Researcher)	1
7 Product Manager	11
8 Scrum Master	3
9 Software Architect	1
10 Software Engineer (Android)	17
11 Software Engineer (Back End)	81
12 Software Engineer (Front End)	44
13 Software Engineer (iOS)	3

Pekerjaan	jumlah_resign	jumlah_belum_resign
0 Data Analyst	8.0	8
1 Data Engineer	3.0	7
2 Product Design (UI & UX)	9.0	15
3 Product Manager	6.0	11
4 Software Engineer (Android)	7.0	17
5 Software Engineer (Back End)	28.0	81
6 Software Engineer (Front End)	28.0	44
7 DevOps Engineer	NaN	3
8 Digital Product Manager	NaN	2
9 Machine Learning Engineer	NaN	2
10 Product Design (UX Researcher)	NaN	1
11 Scrum Master	NaN	3
12 Software Architect	NaN	1
13 Software Engineer (iOS)	NaN	3

Pekerjaan	jumlah_resign
0 Data Analyst	8
1 Data Engineer	3
2 Product Design (UI & UX)	9
3 Product Manager	6
4 Software Engineer (Android)	7
5 Software Engineer (Back End)	28
6 Software Engineer (Front End)	28

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

Resign Reason Analysis for Employee Attrition Management Strategy

- Untuk mendapat persentase resign dari setiap bidang, diperlukan proses agregasi data
 - Mengisi null-value pada kolom yang sesuai,
 - Membuat kolom baru untuk menjumlahkan jumlah karyawan yang resign dan belum resign
 - Menghitung persentase %resign, dari masing-masing bidang pekerjaan

	Pekerjaan	jumlah_resign	jumlah_belum_resign
0	Data Analyst	8.0	8
1	Data Engineer	3.0	7
2	Product Design (UI & UX)	9.0	15
3	Product Manager	6.0	11
4	Software Engineer (Android)	7.0	17
5	Software Engineer (Back End)	28.0	81
6	Software Engineer (Front End)	28.0	44
7	DevOps Engineer	NaN	3
8	Digital Product Manager	NaN	2
9	Machine Learning Engineer	NaN	2
10	Product Design (UX Researcher)	NaN	1
11	Scrum Master	NaN	3
12	Software Architect	NaN	1
13	Software Engineer (iOS)	NaN	3

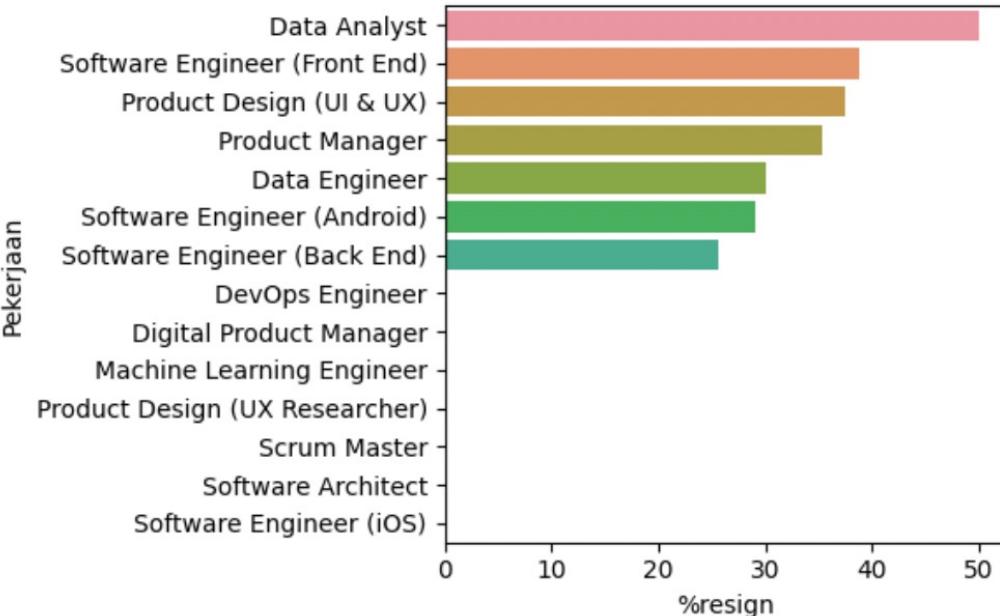


	Pekerjaan	jumlah_resign	jumlah_belum_resign	jumlah_karyawan	%existing	%resign
0	Data Analyst	8.0	8	16.0	50.00	50.00
6	Software Engineer (Front End)	28.0	44	72.0	61.11	38.89
2	Product Design (UI & UX)	9.0	15	24.0	62.50	37.50
3	Product Manager	6.0	11	17.0	64.71	35.29
1	Data Engineer	3.0	7	10.0	70.00	30.00
4	Software Engineer (Android)	7.0	17	24.0	70.83	29.17
5	Software Engineer (Back End)	28.0	81	109.0	74.31	25.69
7	DevOps Engineer	0.0	3	3.0	100.00	0.00
8	Digital Product Manager	0.0	2	2.0	100.00	0.00
9	Machine Learning Engineer	0.0	2	2.0	100.00	0.00
10	Product Design (UX Researcher)	0.0	1	1.0	100.00	0.00
11	Scrum Master	0.0	3	3.0	100.00	0.00
12	Software Architect	0.0	1	1.0	100.00	0.00
13	Software Engineer (iOS)	0.0	3	3.0	100.00	0.00

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

Resign Reason Analysis for Employee Attrition Management Strategy

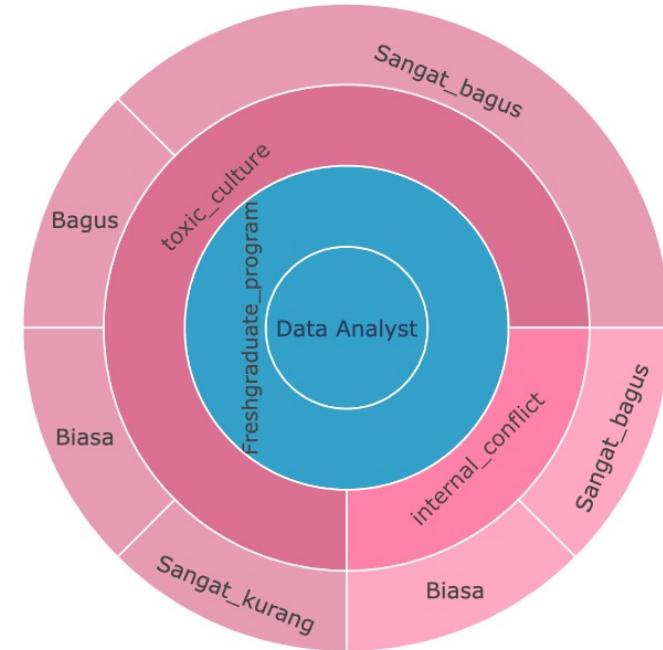
Tingkat Resign Berdasarkan Bidang Pekerjaan



- Tingkat resign yang tertinggi berada pada bidang pekerjaan Data Analyst yakni 50 %, diikuti oleh Software Engineer (Front End) 38,89%, dan Product Design (UI dan UX) 37,50%
- Jumlah karyawan sebagai Data Analyst yang resign sebanyak 8 orang.

Resign Reason Analysis for Employee Attrition Management Strategy

- Dari 8 orang Data Analyst yang resign, kita perlu mengetahui latar belakang mereka berdasarkan Jenjang Karir, Alasan Resign dan Performance selama bekerja.
- Berdasarkan jenjang karir, ke-delapan data analyst yang resign berasal dari jenjang **freshgraduate**.
- Mereka yang resign mayoritas beralasan **toxic culture** dan alasan lainnya seperti **internal conflict**. Kedua alasan tersebut merupakan alasan yang berasal dari kondisi bekerja yang tidak nyaman.
- Lima dari Delapan karyawan memiliki performance pekerjaan dari **bagus** sampai **sangat bagus**. Hal ini cukup disayangkan, karena dari sisi pekerjaan, mereka dapat diandalkan, tetapi memilih untuk resign.
- Untuk Manajemen, perlu melakukan root cause analysis penyebab adanya toxic culture di perusahaan, apakah dari sisi perusahaannya atau sisi perilaku karyawannya.



- Dilakukan Feature Engineering lanjutan:
 - usia_hiring : mengurangkan tahun lahir dan tahun hirring
 - lama_kerja : mengurangkan tahun saat ini atau tahun saat resign dengan tahun hirring
 - JumlahKeikutsertaanProjek : diberi 0 untuk ikut serta, 1 untuk tidak ikut serta
 - KategoriPekerjaan : mengelompokkan pekerjaan dalam bidang software, data, dan product
- Handling Outlier
 - Sebelum handlier outlier : 278 value, sesudah handling outlier : 273 value
- Feature Encoding
 - Frequency Encoding : HiringPlatform
 - Label Encoding : Performance Pegawai, Tingkat Pendidikan, KategoriPekerjaan dan Jenjang karir
 - One Hot Encoding : Status Pernikahan, Jenis Kelamin, Status Kepegawaian, dan Asal Daerah

- Dropping feature (1) :

- TanggalLahir
- TanggalHiring
- TanggalPenilaianKaryawan
- TanggalResign
- tahun_hiring
- tahun_resign
- tahun_lahir
- JumlahKeikutsertaanProjek

- Dropping feature (2) :

- EnterpriseID
- StatusPernikahan
- JenisKelamin
- StatusKepegawaian
- Pekerjaan
- KategoriPekerjaan
- AsalDaerah
- HiringPlatform
- AlasanResign

*Dropping feature (1) dilakukan sebelum encoding,

*Dropping feature (2) dilakukan setelah encoding

Build an Automated Resignation Behavior Prediction using Machine Learning

● Checking Kembali :

- Null Value : tidak ada
- Duplicated Data : tidak ada

```
df_model.duplicated().sum()
```

0

```
df_model.isna().sum()
```

```
JenjangKarir 0  
PerformancePegawai 0  
SkorSurveyEngagement 0  
SkorKepuasanPegawai 0  
JumlahKeterlambatanSebulanTerakhir 0  
JumlahKetidakhadiran 0  
TingkatPendidikan 0  
is_resign 0  
usia_hiring 0  
lama_kerja 0  
JumlahKeikutsertaanProjek_value 0  
platform_hiring_encode 0  
StatusPernikahan_Belum_menikah 0  
StatusPernikahan_Bercerai 0  
StatusPernikahan_Lainnya 0  
StatusPernikahan_Menikah 0  
JenisKelamin_Pria 0  
JenisKelamin_Wanita 0  
StatusKepegawaian_FullTime 0  
StatusKepegawaian_Internship 0  
StatusKepegawaian_Outsource 0  
AsalDaerah_Jakarta_Barat 0  
AsalDaerah_Jakarta_Pusat 0  
AsalDaerah_Jakarta_Selatan 0  
AsalDaerah_Jakarta_Timur 0  
AsalDaerah_Jakarta_Utara 0  
KategoriPekerjaan_Data 0  
KategoriPekerjaan_Product 0  
KategoriPekerjaan_Software 0  
dtype: int64
```

```
df_model.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
Int64Index: 273 entries, 0 to 286  
Data columns (total 29 columns):  
 #   Column           Non-Null Count  Dtype     
 ---  --  
 0   JenjangKarir    273 non-null     int64    
 1   PerformancePegawai 273 non-null     int64    
 2   SkorSurveyEngagement 273 non-null     int64    
 3   SkorKepuasanPegawai 273 non-null     float64  
 4   JumlahKeterlambatanSebulanTerakhir 273 non-null     float64  
 5   JumlahKetidakhadiran 273 non-null     float64  
 6   TingkatPendidikan 273 non-null     int64    
 7   is_resign        273 non-null     int64    
 8   usia_hiring      273 non-null     int64    
 9   lama_kerja       273 non-null     int64    
 10  JumlahKeikutsertaanProjek_value 273 non-null     int64    
 11  platform_hiring_encode 273 non-null     float64  
 12  StatusPernikahan_Belum_menikah 273 non-null     uint8    
 13  StatusPernikahan_Bercerai 273 non-null     uint8    
 14  StatusPernikahan_Lainnya 273 non-null     uint8    
 15  StatusPernikahan_Menikah 273 non-null     uint8    
 16  JenisKelamin_Pria 273 non-null     uint8    
 17  JenisKelamin_Wanita 273 non-null     uint8    
 18  StatusKepegawaian_FullTime 273 non-null     uint8    
 19  StatusKepegawaian_Internship 273 non-null     uint8    
 20  StatusKepegawaian_Outsource 273 non-null     uint8    
 21  AsalDaerah_Jakarta_Barat 273 non-null     uint8    
 22  AsalDaerah_Jakarta_Pusat 273 non-null     uint8    
 23  AsalDaerah_Jakarta_Selatan 273 non-null     uint8    
 24  AsalDaerah_Jakarta_Timur 273 non-null     uint8    
 25  AsalDaerah_Jakarta_Utara 273 non-null     uint8    
 26  KategoriPekerjaan_Data 273 non-null     uint8    
 27  KategoriPekerjaan_Product 273 non-null     uint8    
 28  KategoriPekerjaan_Software 273 non-null     uint8    
dtypes: float64(4), int64(8), uint8(17)  
memory usage: 40.4 KB
```

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

- Split Data Training dan Data Set:

- Perbandingan 8:2
- `y_train` memiliki perbandingan kelas yang tidak seimbang
- Selanjutnya dilakukan handling imbalanced class menggunakan 5 teknik:
 - RandomUndersampler
 - TomekLinks
 - EditedNearestNeighbours
 - SMOTEENN, dan
 - SMOTETomek
- Untuk Modelling Machine Learning menggunakan model XGBoost Classifier
- Untuk menggabungkan Handling Imbalanced Class dan Model Machine Learning menggunakan Pipeline.
- Untuk mengevaluasi Handling Imbalanced Class dan Model Machine Learning menggunakan RepeatedStratifiedKFold, dengan nilai accuracy, precision dan recall score.

```
y_train.value_counts()  
0    150  
1    68  
Name: is_resign, dtype: int64
```

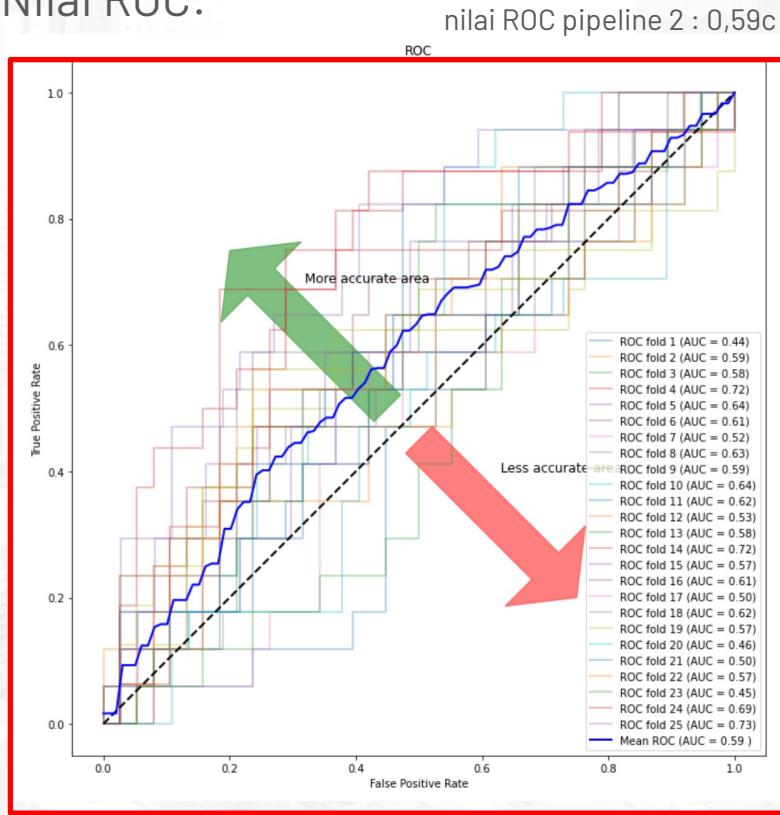
Build an Automated Resignation Behavior Prediction using Machine Learning

- Evaluasi:

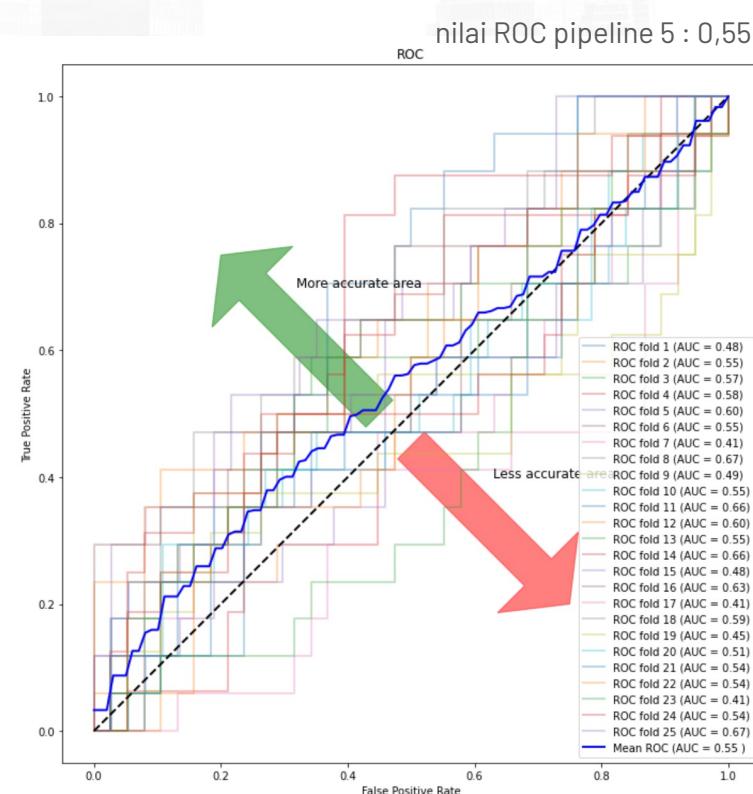
Handling Imbalanced Class	Mean Accuracy	Mean Precision	Mean Recall
RandomUndersampler	0.5490	0.5364	0.5421
TomekLinks	0.6375	0.5589	0.5456
EditedNearestNeighbours	0.5034	0.5270	0.5310
SMOTEENN	0.5254	0.5426	0.5481
SMOTETomek	0.6463	0.5532	0.5414

Build an Automated Resignation Behavior Prediction using Machine Learning

- Nilai ROC:

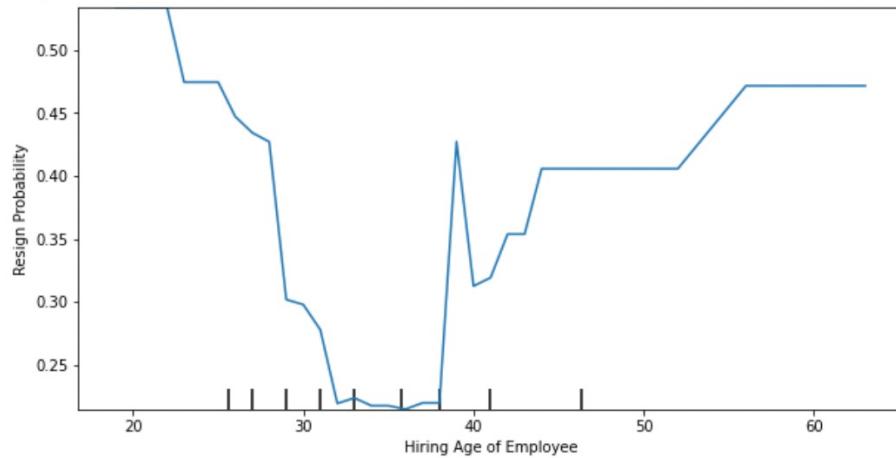


nilai ROC dengan handling imbalance class Tomek Link dan XG Boost lebih baik

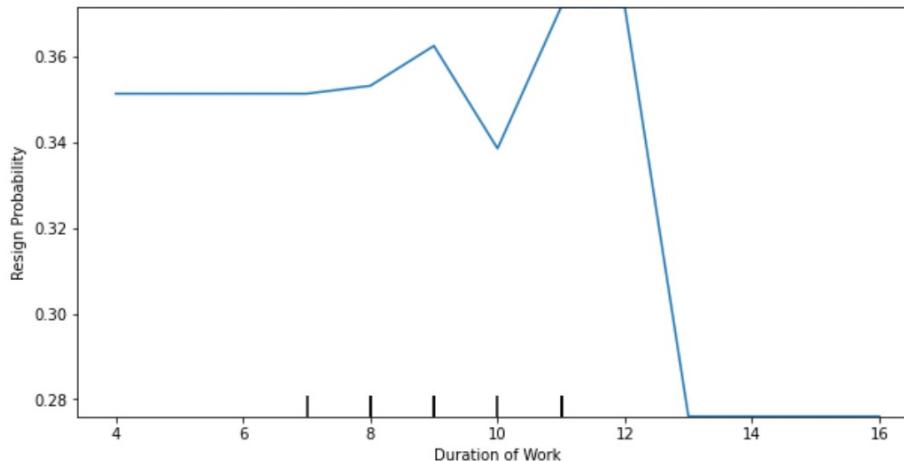


Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

Presenting Machine Learning Products to the Business Users



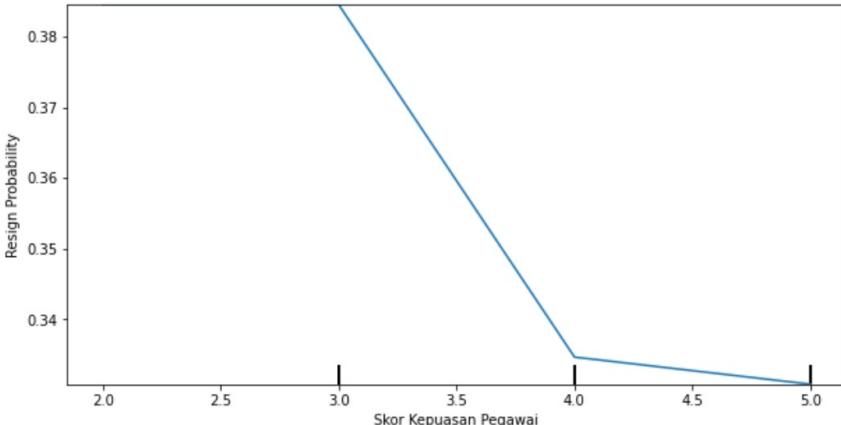
Berdasarkan diagram disamping,
Usia karyawan saat direkrut 20-30 tahun,
merupakan usia dimana tingkat resign yang
tinggi. Sedangkan usia 30-40 tahun,
merupakan usia dengan tingkat resign rendah



Berdasarkan lama bekerja, karyawan yang
sudah bekerja lebih dari 10-12 tahun
cenderung memiliki tingkat resign yang lebih
tinggi dari pada < 10 tahun, dan > 12 tahun.

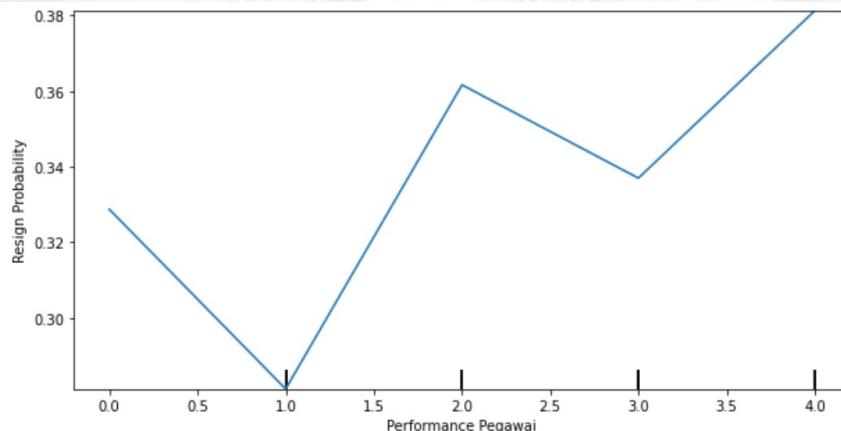
Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

Presenting Machine Learning Products to the Business Users



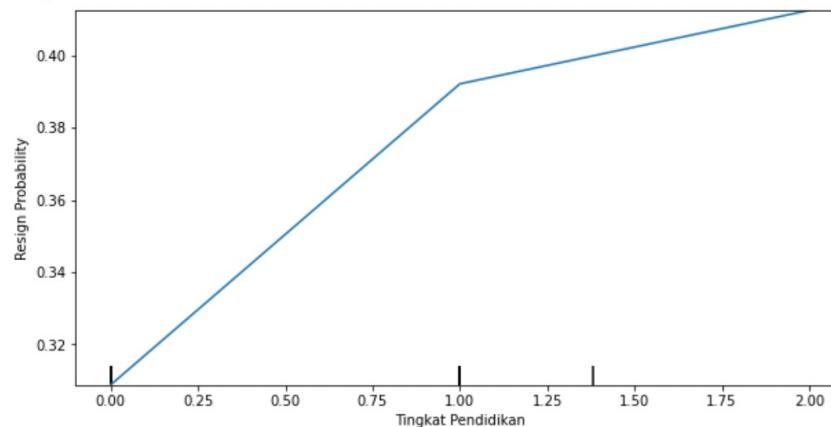
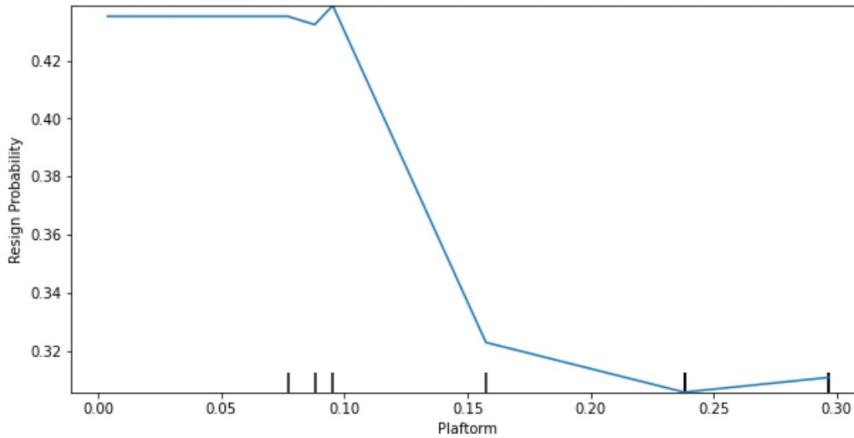
Berdasarkan diagram disamping,

Jika dilihat dari Skor Kepuasan Pegawai dalam bekerja, Karyawan dengan skor rendah, cenderung akan resign daripada karyawan dengan skor kepuasannya tinggi.



Berbanding terbalik dengan Performance Pegawai, Pegawai dengan Performance rendah cenderung tidak resign, dan Pegawai dengan performance baik justru cenderung akan resign.

Presenting Machine Learning Products to the Business Users

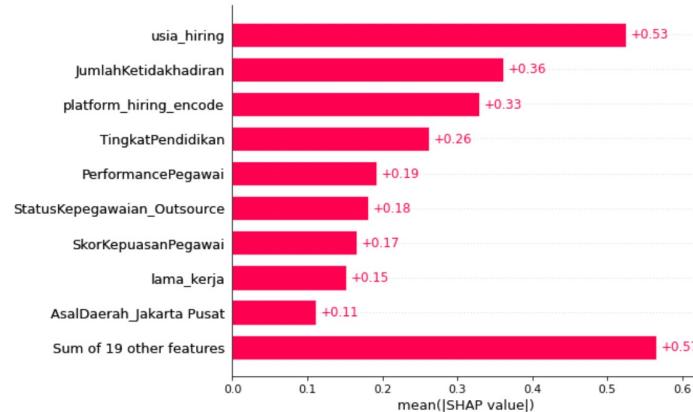


Berdasarkan diagram disamping,

Jika dilihat dari Plaftorm Rekuriter Karyawan, Platform yang berasal dari Online Web application, Other, dan Website memiliki tingkat resign yang tinggi (dilihat dari jumlah frekuensinya yang sedikit)

Berdasarkan tingkat Pendidikan, semakin tinggi tingkat Pendidikan, semakin tinggi tingkat kemungkinan untuk resignnya. Untuk tingkat Pendidikan Sarjana, tingkat resign cendrung rendah dibandingkan dengan tingkat Pendidikan Doktoral.

Presenting Machine Learning Products to the Business Users



Berdasarkan tingkat resign, variabel yang berpengaruh adalah usia saat hiring, jumlah ketidak hadiran, dan platform hiring.

Sehingga, beberapa rekomendasi yang perlu dipertimbangkan :

- Usia hiring (perlu melakukan recruitment pada usia karyawan antara 30-40 tahun).
- Jika perlu merekrut usia 20-30 tahun, perlu memperhatikan lingkungan kerja, karena alasan resign yang terbesar adalah lingkungan kerja yang tidak nyaman. Pihak managemen perlu melihat Kembali penyebab ketidaknyamanannya lingkungan kerja.
- Menghindari rekruitmen via website langsung. Perlu melakukan recruitment via LinkedIn, dan Indeed. Karena para job seeker lebih menyukai apply via sarana tersebut.