

Improving Employee Retention by Predicting Employee Attrition Using Machine Learning

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



Created by:

Sendhy Boedhi Satriya
sendhyboedhisatriya@gmail.com

<https://www.linkedin.com/in/masendhy/>

About_ Hi! I'm Sendhy, an ambitious and eager to learn deeply interested in Data and Tech related role such as Data Science currently, I am focused on developing my skills and experience in data science, developing my science knowledge, as well as developing some of the soft skills needed in the future.

Job Experience_

Web Developer – **Remote Worker Indonesia** (Present)
Supervisor – **PT.Bank Rakyat Indonesia,Tbk** (2013 – 2018)

Skills_

Skilled in Data Science and Web Programming using Python, SQL, Tableau, Power BI, Django, HTML, CSS, Javascript (React JS, NodeJS), PHP (Laravel), Streamlit, Webflow, and MySQL.

Connect_

<https://medium.com/@masendhy>
<https://github.com/masendhy>

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

- Melakukan proses EDA (melakukan pemeriksaan terhadap tipe data dari setiap kolom, jumlah baris dan kolom, identifikasi missing values, analisa statistic deskriptif, deteksi outliers, analisa distribusi dan analisa korelasi, melakukan pengelompokan fitur data berdasarkan tipe datanya)
- Melakukan Data Processing :
 - membuat kolom baru yaitu kolom *Age* : untuk mengetahui umur karyawan yang melakukan resign.

```
● ● ●

# transform column TanggalLahir
X['TanggalLahir'] = pd.to_datetime(X['TanggalLahir'],
format='%Y/%m/%d')
}

# create Age column
X['Age'] = 2023 - X['TanggalLahir'].dt.year
```

- menghapus kolom yang tidak relevan : untuk menyederhanakan dataset dan fokus pada fitur-fitur yang lebih penting.

```
● ● ●

#drop unneeded columns
X =
df.drop(['EnterpriseID', 'JumlahKeikutsertaanProjek', 'JumlahKeterlambatanSebulanTerakhir', 'JumlahKetidakhadiran', 'IkutProgramLOP', 'TanggalResign', 'TanggalPenilaianKaryawan', 'TanggalHiring', 'AlasanResign', 'PernahBekerja', 'Email', 'NomorHP', 'JumlahKetidakhadiran', 'JumlahKeterlambatanSebulanTerakhir', 'JumlahKeikutsertaanProjek', 'HiringPlatform', 'AsalDaerah', 'Username'], axis=1)
```