

zenius



Kampus
Merdeka
INDONESIA JAYA

Final Project Presentation

Nomor Kelompok: 2

Nama Mentor: Aditya Bariq

Nama:

- Nirma Oktaviani
- Nur Alifiah

Machine Learning Class

Program Studi Independen Bersertifikat
Zenius Bersama Kampus Merdeka



Petunjuk

- Waktu presentasi adalah 5 menit (tentatif, tergantung dari banyaknya kelompok yang mendaftarkan diri)
- Waktu tanya jawab adalah 5 menit
- Silakan menambahkan gambar/visualisasi pada slide presentasi
- Upayakan agar tetap dalam format poin-poin (ingat, ini presentasi, bukan esai)
- Jangan masukkan *code* ke dalam slide presentasi (tidak usah memasukan screenshot jupyter notebook)

1. Latar Belakang
2. Explorasi Data dan Visualisasi
3. Modelling
4. Kesimpulan

Latar Belakang

Latar Belakang Project

Sumber Data:

<https://www.kaggle.com/datasets/barun2104/telecom-churn?datasetId=567482>

Problem: **classification**

Tujuan:

- Mengetahui faktor-faktor yang mempengaruhi *Churn*
- Melakukan classification untuk mengetahui model terbaik dengan melihat tingkat akurasi yang didapatkan

Explorasi Data dan Visualisasi

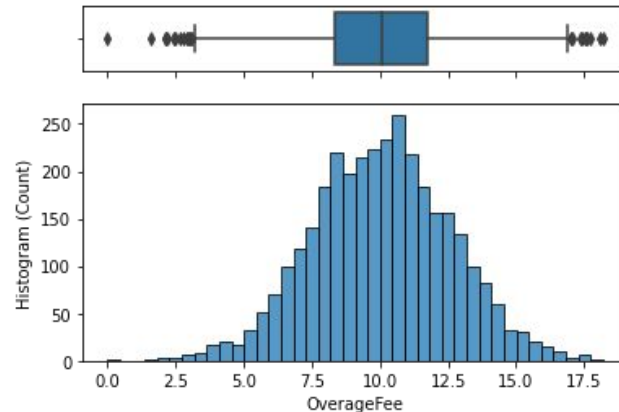
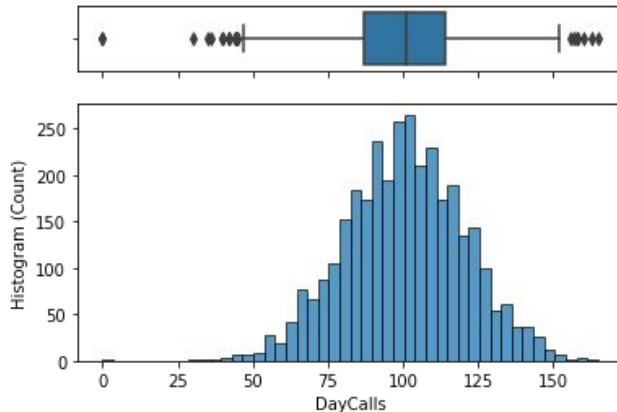
Business Understanding

Churn merupakan istilah yang digunakan untuk pelanggan yang telah menghentikan layanan provider tertentu atau pindah ke layanan provider lain. Dalam keberjalanannya perusahaan terus meminimalisir churn agar tidak kehilangan pelanggan dengan mengeluarkan berbagai penawaran menarik. Hal ini dikarenakan biaya yang dibutuhkan untuk menarik pelanggan baru lebih besar dibandingkan mempertahankan pelanggan. Churn sangat penting dalam bisnis untuk menilai apakah perusahaan sudah menjawab kebutuhan masyarakat atau tidak.

Data Cleansing

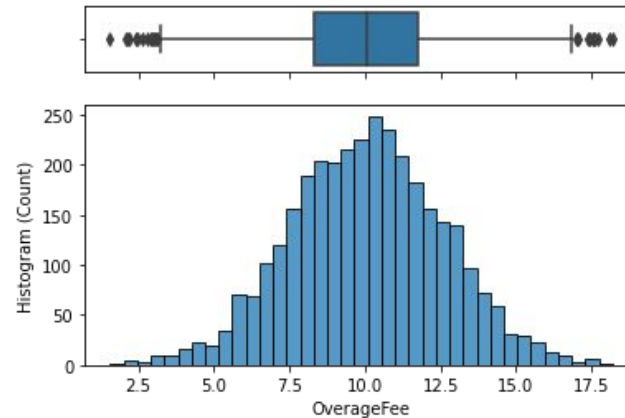
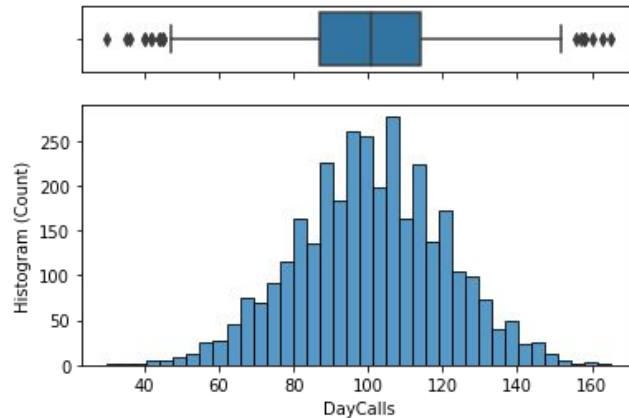
Data yang diperoleh tidak memiliki *missing value*, namun terdapat beberapa kejanggalan pada data, yaitu:

- *Outliers* pada kolom 'DayCalls', terdapat 2 pelanggan yang memiliki nilai rata-rata telepon harian 0, namun sudah berlangganan sudah berlangganan selama beberapa minggu dan membayar sewa bulanan.
- *Outliers* pada kolom 'Overage Fee'. terdapat 1 pelanggan yang memiliki biaya langganan terbesar yang dikeluarkan selama berlangganan = 0, padahal customer sudah berlangganan selama beberapa minggu dan membayar sewa bulanan



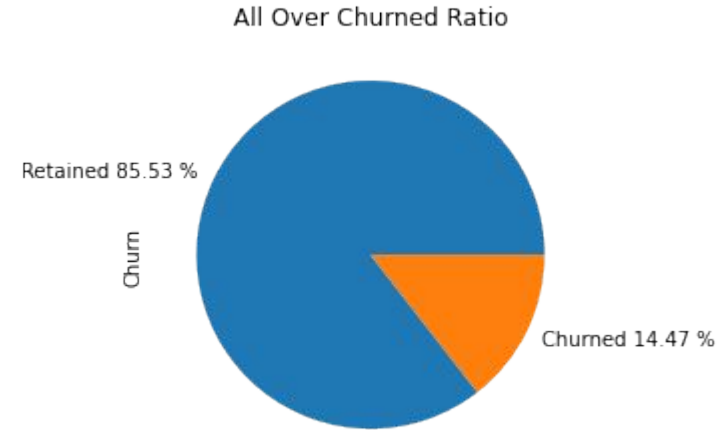
Data Cleansing

Untuk mengatasi kegagalan yang ada, maka data yang mengandung *outliers* akan dihilangkan.



Exploratory Data Analysis

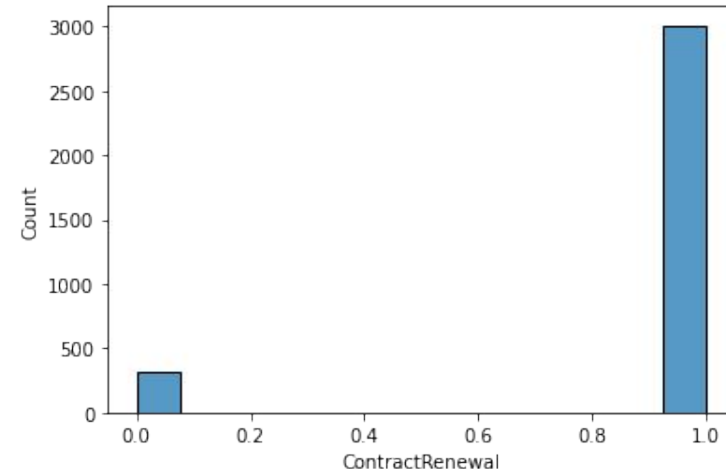
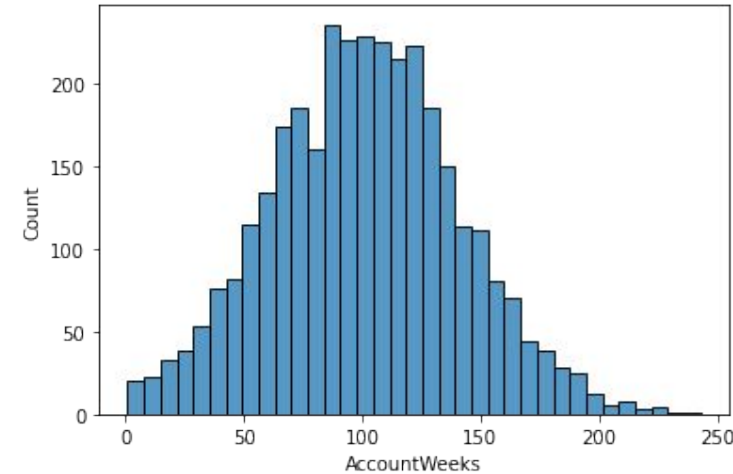
Berdasarkan gambar dapat dilihat bahwa ratio pelanggan yang membatalkan layanan lebih kecil dibandingkan dengan yang masih memakai layanan telekomunikasi. oleh karena itu, perusahaan masih bisa mempertahankan pelanggannya



Exploratory Data Analysis

Berdasarkan histogram di samping, dapat diketahui bahwa:

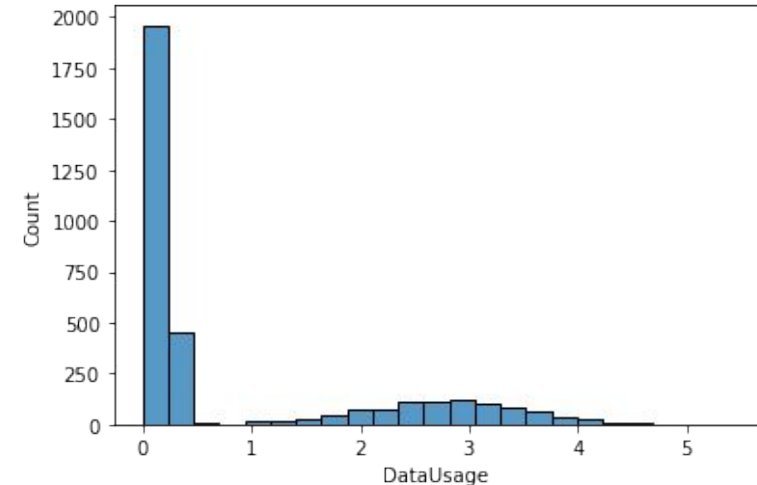
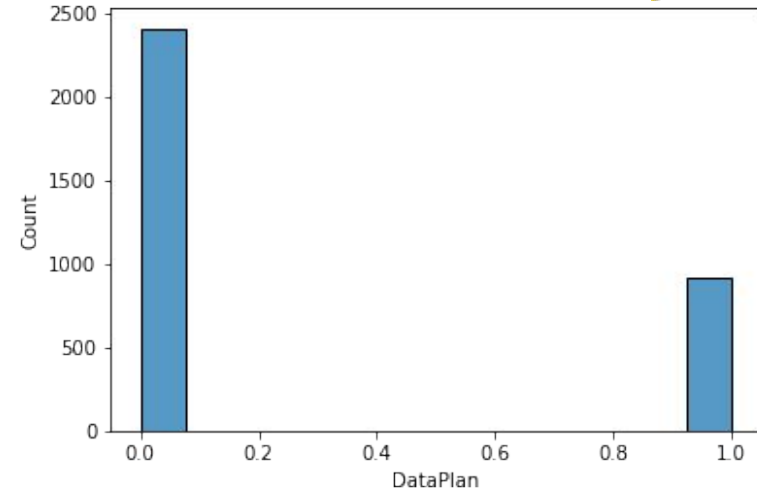
- Pada histogram Account Weeks dapat dilihat bahwa rata-rata lama pelanggan berlangganan berada pada kisaran 0-200 minggu
- Pada histogram Contract Renewal dapat dilihat bahwa lebih banyak pelanggan yang memperbarui kontrak untuk berlangganan dibandingkan dengan yang tidak



Exploratory Data Analysis

Berdasarkan histogram di samping, dapat diketahui bahwa:

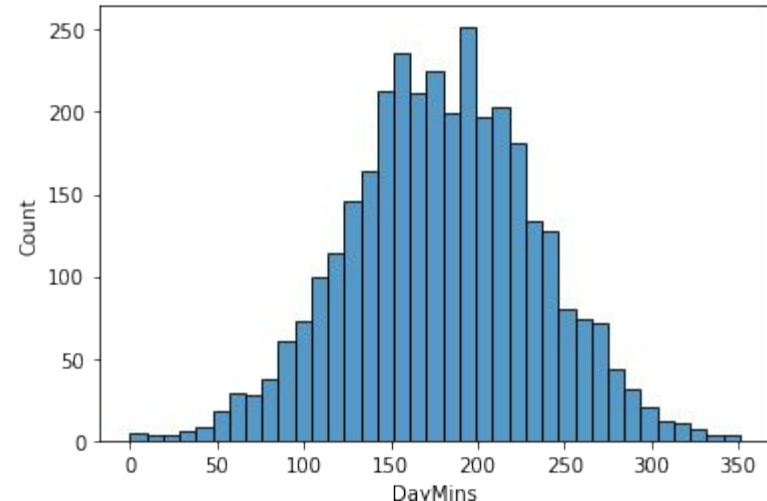
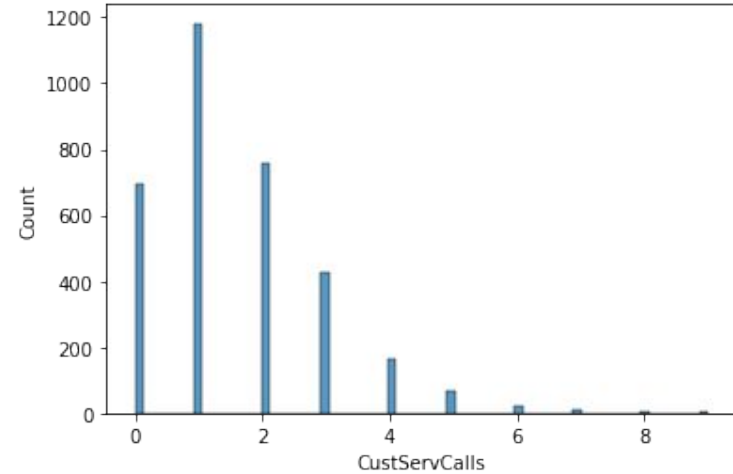
- Pada histogram Data Plan dapat dilihat bahwa lebih banyak pelanggan yang tidak memiliki paket data dibandingkan yang memiliki paket data
- Pada histogram Data Usage dapat dilihat bahwa mayoritas pelanggan tidak memakai data sama sekali atau hanya sebesar 1 gigabytes setiap bulannya



Exploratory Data Analysis

Berdasarkan histogram di samping, dapat diketahui bahwa:

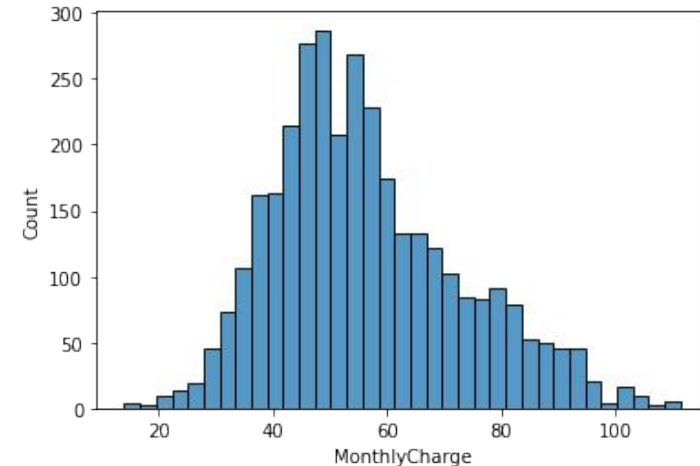
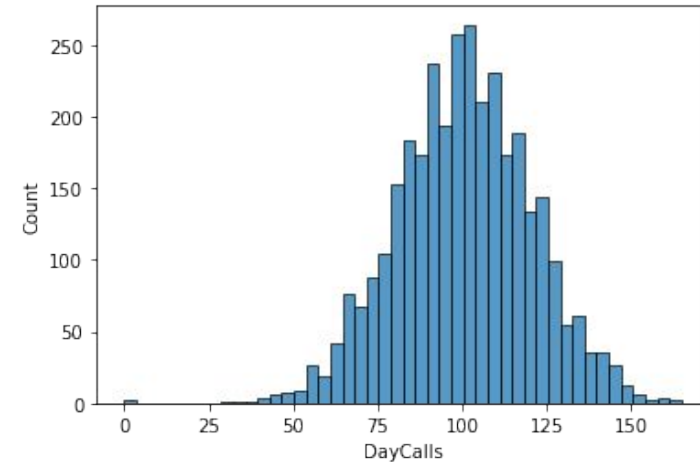
- Pada histogram CustServCalls dapat dilihat bahwa pelanggan yang paling banyak melakukan panggilan Customer Service yaitu dalam 1 kali panggilan
- Pada histogram DayMins dapat dilihat bahwa rata-rata lama pelanggan menggunakan layanan berada pada kisaran 50-300 menit perbulannya



Exploratory Data Analysis

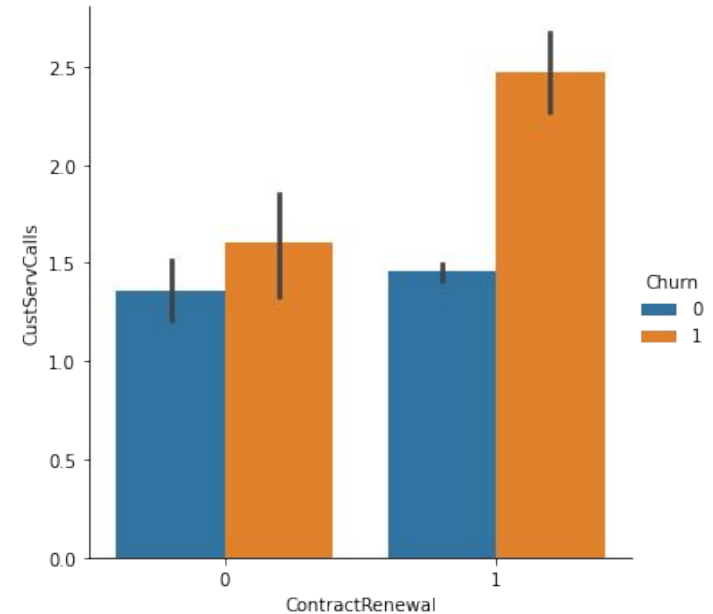
Berdasarkan histogram di samping, dapat diketahui bahwa:

- Pada histogram Day Calls dapat dilihat bahwa rata-rata lama pelanggan menggunakan layanan perharinya berada pada kisaran 50-150
- Pada histogram Monthly Charge dapat dilihat bahwa rata-rata pelanggan membayar biaya berlangganan sebesar 20-100 dollar per bulannya



Exploratory Data Analysis

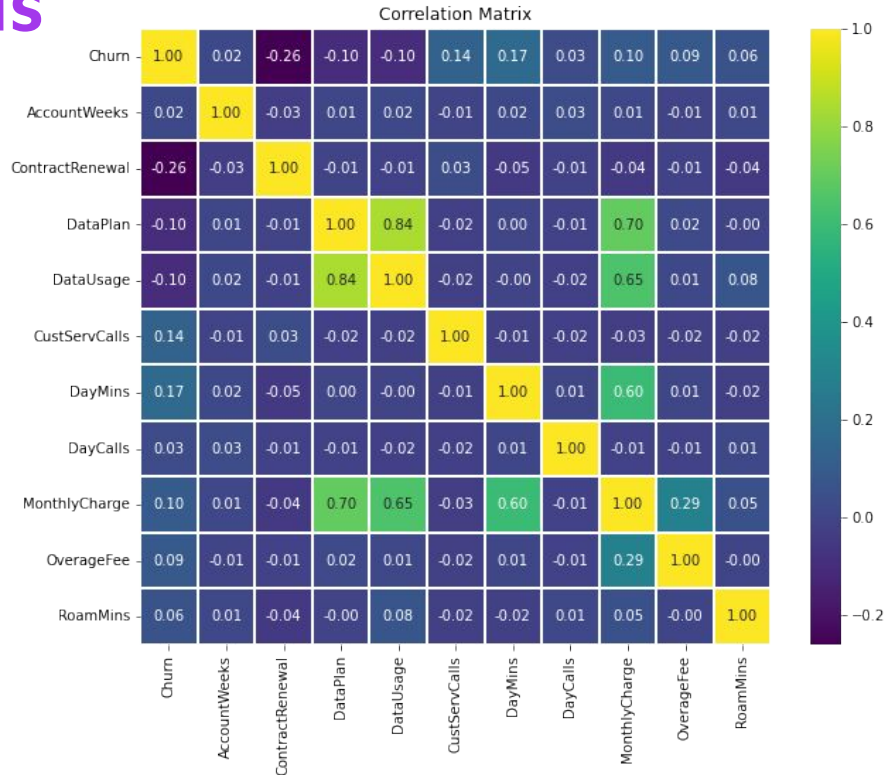
Berdasarkan data dapat dilihat bahwa semakin sering pelanggan menghubungi Customer Service maka besar kemungkinan pelanggan untuk churn



Exploratory Data Analysis

Korelasi antar variabel dapat dikatakan kuat apabila bernilai positif dan tinggi mendekati angka 1. Korelasi dengan menggunakan metode spearman dapat dilihat bahwa korelasi yang paling kuat tiap variabel, yaitu:

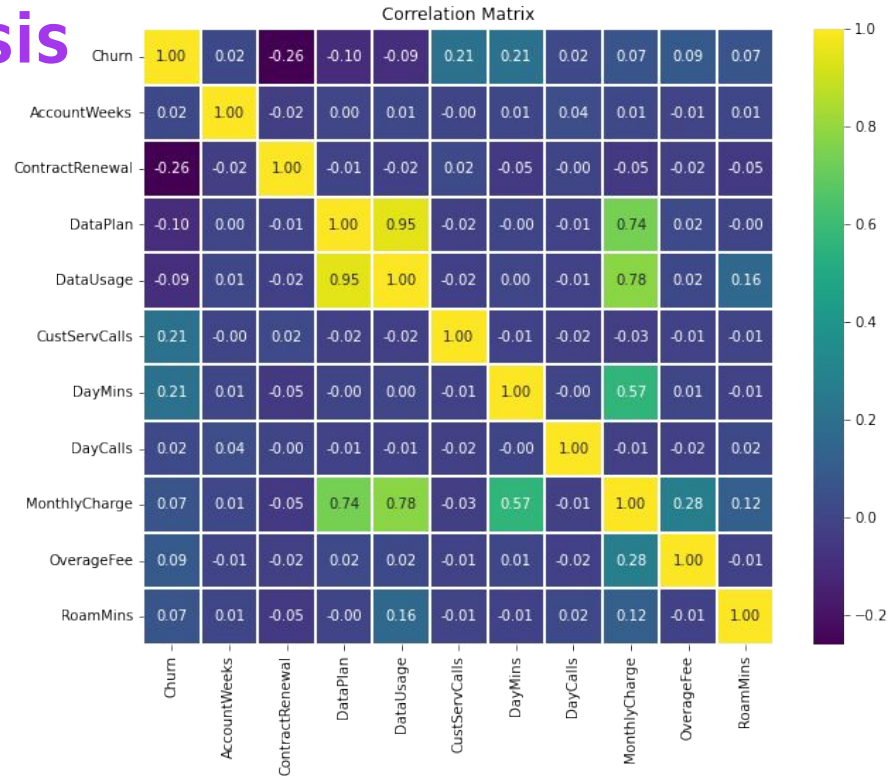
- Data Plan dengan Data Usage
- Data Plan dengan Monthly Charge
- Data Usage dengan Monthly Charge
- Day Mins dengan Monthly Charge



Exploratory Data Analysis

Korelasi dengan menggunakan metode Pearson dapat dilihat bahwa korelasi yang paling kuat tiap variabel, yaitu:

- Data Plan dengan Data Usage
- Data Usage dengan Monthly Charge
- Data Plan dengan Monthly Charge
- Day Mins dengan Monthly Charge



Modelling



Evaluation Model

Dengan menggunakan 80% data train dan 20% data test menghasilkan:

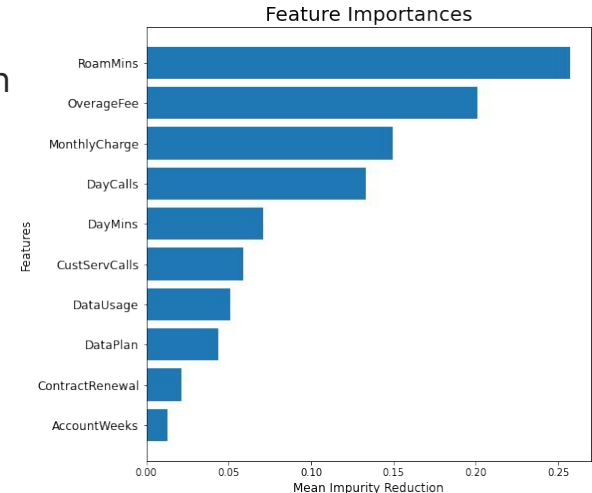
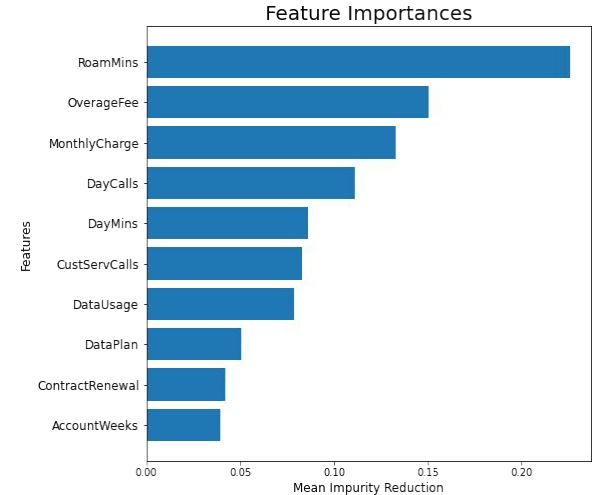
Model/Metode	AUC	f1-score	Accuracy
Decision Tree	0.81	0.95	0.91
Random Forest	0.82	0.97	0.95
Random Forest with <i>hypertuning</i>	0.85	0.97	0.95

Berdasarkan tabel dapat disimpulkan bahwa metode random forest dengan hypertuning merupakan metode terbaik untuk mengklasifikasi data churn.

Conclusion

Perbandingan Metode Decision Tree dan Random Forest dengan Hypertuning untuk Mengklasifikasi Data Churn

- Berdasarkan grafik dapat disimpulkan bahwa data RoamMin merupakan data yang paling mempengaruhi data Churn, diikuti dengan OverageFee dan MonthlyCharge. Maka untuk mengurangi tingkat churn perusahaan bisa lebih memperhatikan pelayanan untuk roaming ataupun memberikan tingkat diskon kepada pelanggan. Selain itu RoamMins juga menjadi faktor utama untuk membedakan customer yang churn dan tidak churn
- Berdasarkan analisis dapat dilihat bahwa semakin sering pelanggan menghubungi Customer Service maka besar kemungkinan pelanggan untuk churn



Terima kasih!

Ada pertanyaan?

zenius



Kampus
Merdeka
INDONESIA JAYA