

zenius

Kampus  
Merdeka  
INDONESIA JAYA

# Final Project Presentation

Nomor Kelompok: 06

Nama Mentor: Rachmadino Noval L.

Nama:

- Anjani Maulaya Nunqaf
- Destiwi Ramadani

Machine Learning Class

Program Studi Independen Bersertifikat  
Zenius Bersama Kampus Merdeka



- 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:

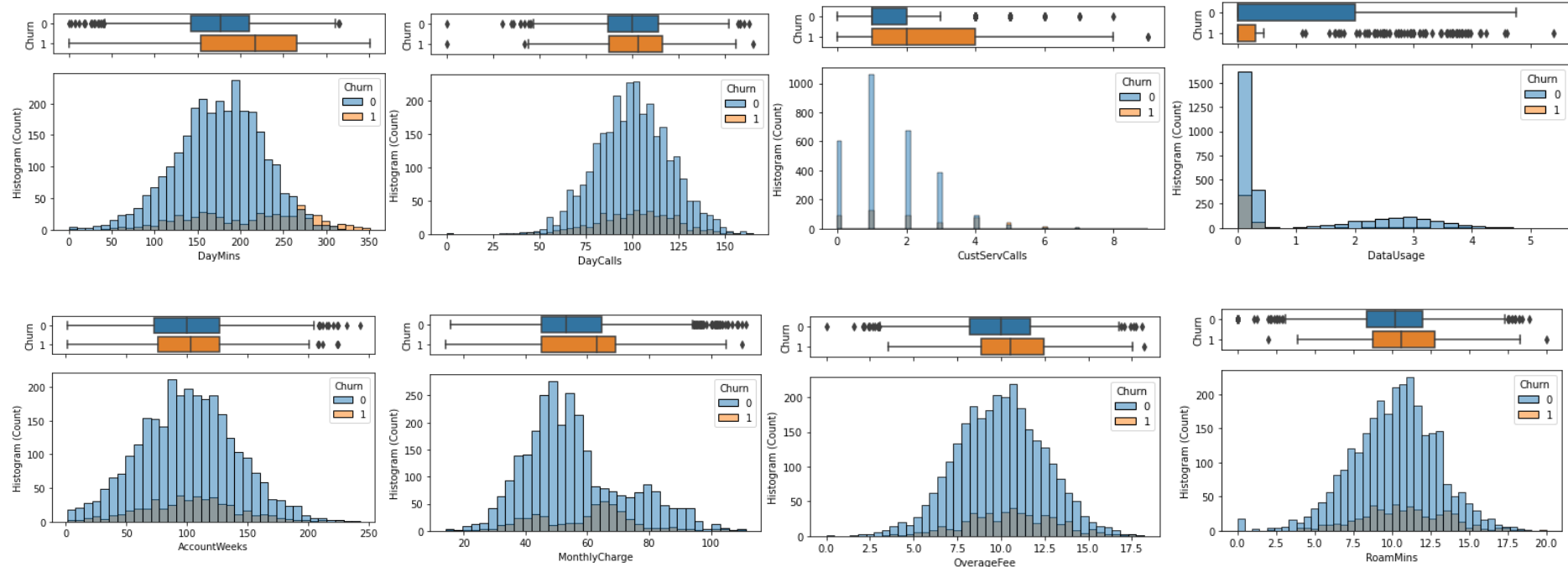
- Memprediksi customer sebuah perusahaan telecom yang mengalami churn
- Memprediksi faktor yang mempengaruhi customer churn
- Mengetahui model/ machine learning yang tepat untuk memprediksi Churn Customer

# Explorasi Data dan Visualisasi

# Business Understanding

- Customer Churn adalah metrik bisnis yang mengukur jumlah pelanggan yang telah berhenti menggunakan produk atau layanan dari suatu perusahaan.
- Customer Churn harus diantisipasi agar perusahaan tidak mengalami kerugian akibat kehilangan pelanggan.
- Kehilangan Customer juga akan mempengaruhi growth rate dari perusahaan

# Data Cleansing



# Data Cleansing

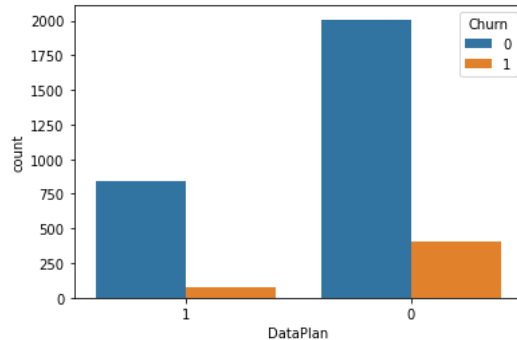
Customer Churn dataset terdiri dari 3333 baris dan 11 kolom

Pada Customer Churn dataset tidak perlu dilakukan Data Cleansing, karena outliernya alami.

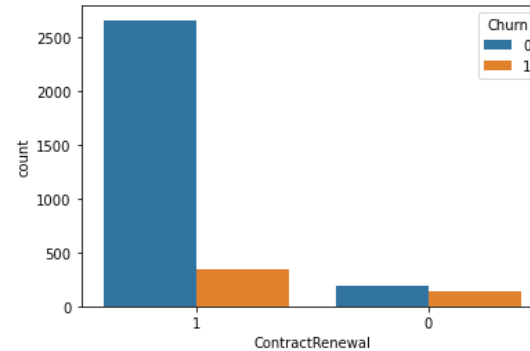


# Exploratory Data Analysis

- Memisahkan data yang *Churn* dan *No Churn* , kemudian membuat visualisasi dari masing-masing data. Membuat grafik dari field data kategori dan numerik.



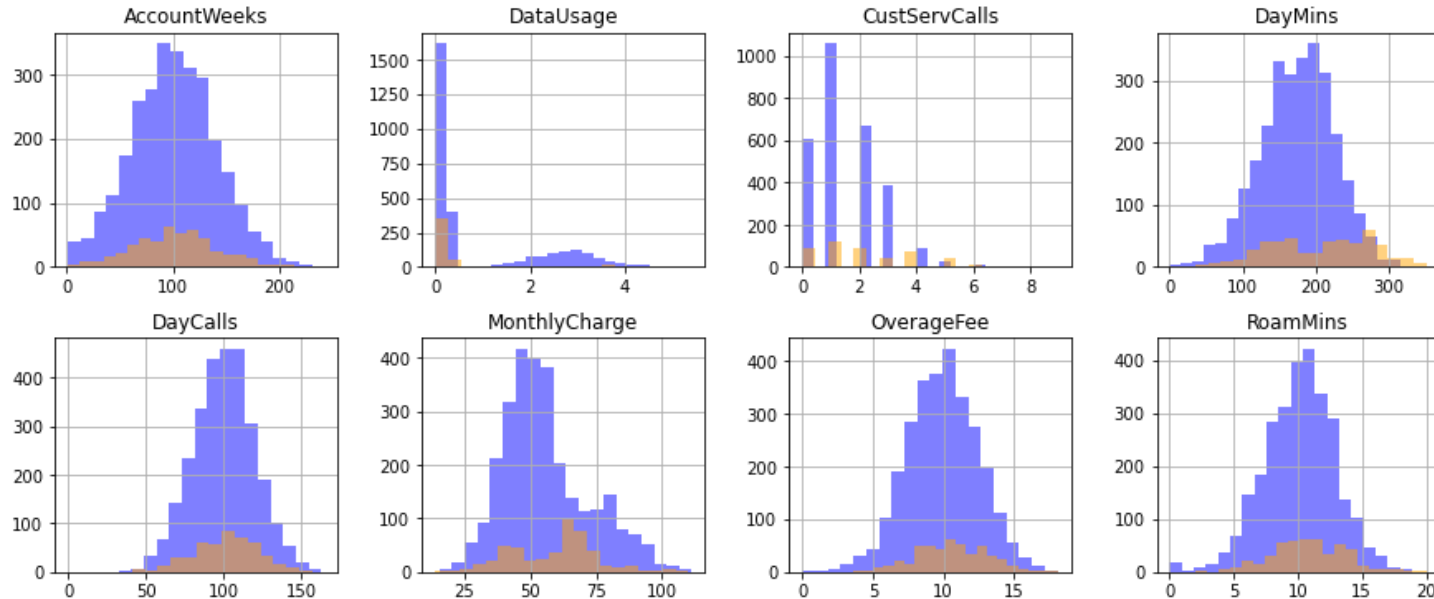
Semakin tidak memiliki data plan  
maka semakin tinggi  
kemungkinan untuk Churn



Pada contract renewal Perbedaan tidak  
signifikan sehingga tidak terlalu  
mempengaruhi Churn

# Exploratory Data Analysis

Plot untuk data numerik



# Exploratory Data Analysis

	Churn	AccountWeeks	ContractRenewal	DataPlan	DataUsage	CustServCalls	DayMins	DayCalls	MonthlyCharge	OverageFee	RoamMins
Churn	1.000000	0.015583	-0.259852	-0.102148	-0.102080	0.136657	0.170677	0.026311	0.094655	0.088582	0.060850
AccountWeeks	0.015583	1.000000	-0.029538	0.008741	0.016805	-0.005942	0.017884	0.032690	0.011784	-0.007956	0.014761
ContractRenewal	-0.259852	-0.029538	1.000000	-0.006006	-0.010479	0.030328	-0.049604	-0.013043	-0.044644	-0.010400	-0.043836
DataPlan	-0.102148	0.008741	-0.006006	1.000000	0.841275	-0.022156	0.002989	-0.012653	0.695608	0.022762	-0.004484
DataUsage	-0.102080	0.016805	-0.010479	0.841275	1.000000	-0.017852	-0.004309	-0.019481	0.646449	0.011492	0.076597
CustServCalls	0.136657	-0.005942	0.030328	-0.022156	-0.017852	1.000000	-0.015032	-0.020957	-0.025401	-0.017801	-0.017374
DayMins	0.170677	0.017884	-0.049604	0.002989	-0.004309	-0.015032	1.000000	0.009391	0.598887	0.006401	-0.016110
DayCalls	0.026311	0.032690	-0.013043	-0.012653	-0.019481	-0.020957	0.009391	1.000000	-0.009012	-0.014064	0.015139
MonthlyCharge	0.094655	0.011784	-0.044644	0.695608	0.646449	-0.025401	0.598887	-0.009012	1.000000	0.289230	0.055563
OverageFee	0.088582	-0.007956	-0.010400	0.022762	0.011492	-0.017801	0.006401	-0.014064	0.289230	1.000000	-0.003210
RoamMins	0.060850	0.014761	-0.043836	-0.004484	0.076597	-0.017374	-0.016110	0.015139	0.055563	-0.003210	1.000000

# Modelling



# Classification Customer Churn

Pada tahap pemodelan kami memecah 2 dataset sebagai :

- Data Train sebesar 60% dari total
- Data Test sebesar 40% dari total

Dalam proses modelling kami menggunakan 3 model classification :

1. Logistic Regression
2. Decision Tree
3. Random Forest

Classifier	Accuracy	Precision	Recall
Logistic Regression	84%	64%	56%
Decision Tree	86%	72%	74%
Random Forest	93%	91%	81%

Model performance paling bagus adalah Random Forest

Model akhir yang kami gunakan dan kami rekomendasikan adalah Random Forest Tree.

Kami membandingkan dengan 3 model yang kami gunakan, sehingga diperoleh Random Forest Tree yang terbaik dengan akurasi 0.93, Recall 0.81 dan Precision 0.91

	precision	recall	f1-score	support
0	0.94	0.98	0.96	1136
1	0.88	0.64	0.74	198
accuracy			0.93	1334
macro avg	0.91	0.81	0.85	1334
weighted avg	0.93	0.93	0.93	1334

# Conclusion



Kemungkinan besar customer churn dikarenakan biaya perbulan dan biaya tambahan terlalu mahal untuk customer atau ada yang lebih murah dari yang ditawarkan oleh perusahaan telekom lainnya, maka dari itu rekomendasi untuk perusahaan agar customer churn menurun adalah dengan memberi diskon atau gift kepada pelanggan yang telah menggunakan rata-rata 100 menit/hari.

**Terima  
kasih!**  
**Ada pertanyaan?**

**zenius**



**Kampus  
Merdeka**  
INDONESIA JAYA

