Churn Predict Analyst 2022

Andrianto Saputra

Data Source

	user_id	gender	age	source	quantity	created_at	price	browser	event_type
0	2021	M	64	Adwords	1	2022-01-01 00:41:00.000000 UTC	39.00	Chrome	home
1	2021	M	64	Email	1	2022-01-01 00:41:00.000000 UTC	39.00	Safari	purchase
2	2021	M	64	Email	1	2022-01-01 00:41:00.000000 UTC	39.00	Safari	department
3	2021	M	64	Email	1	2022-01-01 00:41:00.000000 UTC	39.00	Firefox	purchase
4	2021	M	64	Email	1	2022-01-01 00:41:00.000000 UTC	39.00	Safari	cart

Data source from ip xx.xx.xx.xx
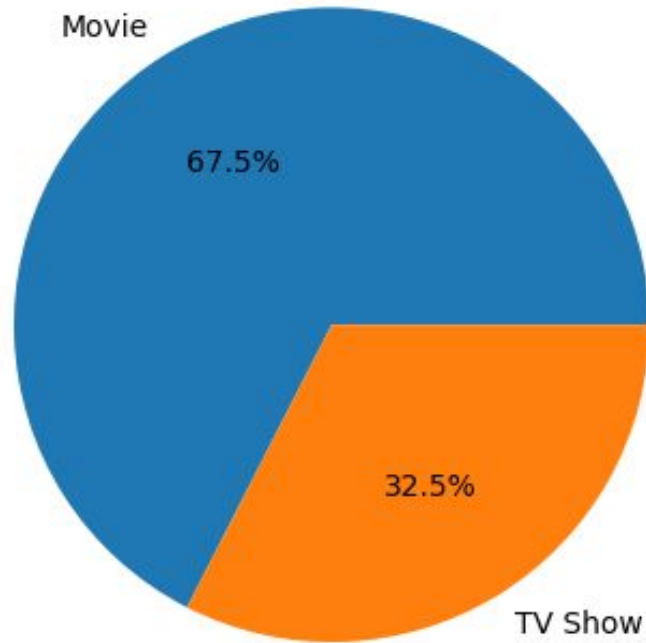
Data source from ip xx.xx.xx.xx

Event Type Distribution

Browser Distribution

Movie Distribution

Dari data netflix tahun 2019 Movie lebih mendominasi sebesar **67.5%** dibandingkan TV show yang hanya memiliki 32.5%

Data Defined Churn

Period	Total
December	3512
November	3289
October	3235
September	2840
August	2883
July	2746
June	2485
May	2267
April	2224
March	2191
February	1966
January	2144

Sampling 2

id	recency	frequency	Money
1	**287**	1	65.000
2	**107**	2	19.000

Churn IF **180 Days** not return

Churn	Total (%)
Yes	56 %
No	44 %

Churn Prediction Model

Churn	Total (%)
Yes	56 %
No	44 %

Train 80%, Testing 20%

Metode
XGB Boost
Decission Tree
SVM

Kesimpulan

- Dari model yang ada akan dipakai beberapa parameter sebagai fitur machine learning churn prediction ,yaitu:
  1.gender
  2.Age
  3.Source
  4.Quantity
  5.Browser
  6.Event type
  7. Price
- Berikut performa metode machine learning yang akan dipakai dengan test periode 2021-2023 sebagai berikut :

Metode	Accuracy		
	2021	2022	2023
XGB Boost / mean	73 %	70 %	75 %
Decission Tree / median	86 %	89 %	75 %