KOBOLD.AI

Customer Churn Prediction

Customer retention is one of the most effective ways to assure continuous revenue for a company. Yet, it is often hard to understand why customers do not return to buy or cancel their contracts. And due to limited resources, it is impossible to try to keep all customers involved and personally contact them.
Thus, one of the approaches in data-driven companies is to try to understand which customers might churn next and give these personalised attention. The idea behind such a prioritisation is to focus capacity on those customers who are most likely to leave. The subsequent activities might be personalised emails, training offerings, coupons or specialised contracts.
As such, customer churn prediction is a way to increase customer satisfaction on the one hand, but also increase customer retention on the other hand.

Meta-Data

Algorithm:	Classification
Start:	29-May-2021 (21:39:29.251419)
Filename:	telco_churn.csv
Dimensions:	(7043, 21)

Pre-Processing

Dropped column 'customerID' due to high dimensionality.

KOBOLD.AI

Input data overview

Attribute	Count	Mean	Range	Mode	Unique
gender	7032	-	-	Male	2
SeniorCitizen	7032	0.16	0.0 - 1.0	0.0	2
Partner	7032	-	-	No	2
Dependents	7032	-	-	No	2
tenure	7032	32.42	1.0 - 72.0	1.0	72
PhoneService	7032	-	-	Yes	2
MultipleLines	7032	-	-	No	3
InternetService	7032	-	-	Fiber optic	3
OnlineSecurity	7032	-	-	No	3
OnlineBackup	7032	-	-	No	3
DeviceProtection	7032	-	-	No	3
TechSupport	7032	-	-	No	3
StreamingTV	7032	-	-	No	3
StreamingMovies	7032	-	-	No	3
Contract	7032	-	-	Month-to-month	3
PaperlessBilling	7032	-	-	Yes	2
PaymentMethod	7032	-	-	Electronic check	4
MonthlyCharges	7032	64.8	18.25 - 118.75	20.05	1584
TotalCharges	7032	2283.3	18.8 - 8684.8	20.2	6530
Churn	7032	-	-	No	2

KOBOLD.AI

Results

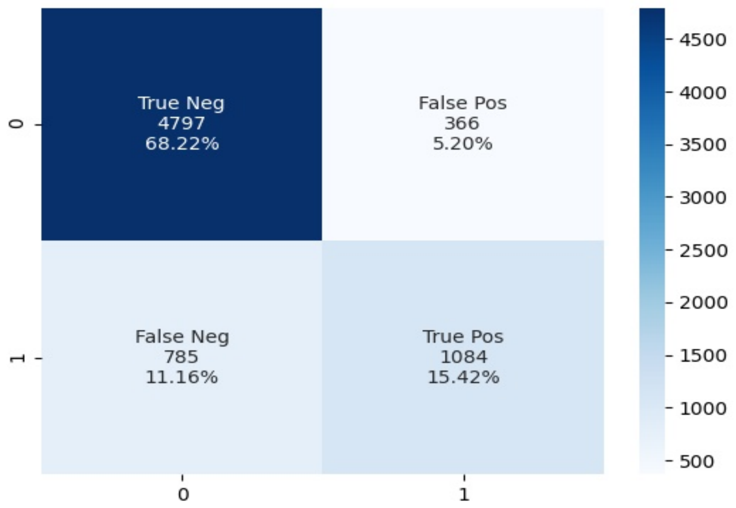
Training accuracy: 94.01%

During training, we train hundreds of statistical models to fit your data as good as possible. The training accuracy score denotes how well the model fits the target variable by comparing the real to the predicted target value.

Testing accuracy: 83.79%

The higher the testing accuracy, the better it will perform in predicting new, unseen examples (generalisation). While highly individual on the input data, everything above 80% is a very good model.

Confusion Matrix



A confusion matrix shows how many predictions matched the real values and which got predicted incorrectly.

KOBOLD.AI

Examples

Correctly classified

gender	Female	Female	Female	Male	Male
SeniorCitizen	1.0	0.0	0.0	0.0	0.0
Partner	No	Yes	Yes	No	No
Dependents	No	No	Yes	No	No
tenure	32.0	72.0	72.0	19.0	36.0
PhoneService	Yes	Yes	No	Yes	Yes
MultipleLines	No	Yes	No phone service	Yes	Yes
InternetService	Fiber optic	Fiber optic	DSL	DSL	Fiber optic
OnlineSecurity	No	Yes	Yes	No	No
OnlineBackup	No	Yes	No	No	No
DeviceProtection	No	Yes	Yes	Yes	No
TechSupport	No	Yes	No	Yes	No
StreamingTV	Yes	Yes	Yes	No	No
StreamingMovies	Yes	Yes	Yes	No	Yes
Contract	Month-to-month	Two year	Two year	Month-to-month	Month-to-month
PaperlessBilling	No	Yes	Yes	No	Yes
PaymentMethod	Electronic check	Bank transfer (automatic)	Credit card (automatic)	Electronic check	Electronic check
MonthlyCharges	90.95	115.5	53.8	59.55	84.75
TotalCharges	2897.95	8312.75	3952.45	1144.6	3050.15
Churn	No	No	No	No	Yes

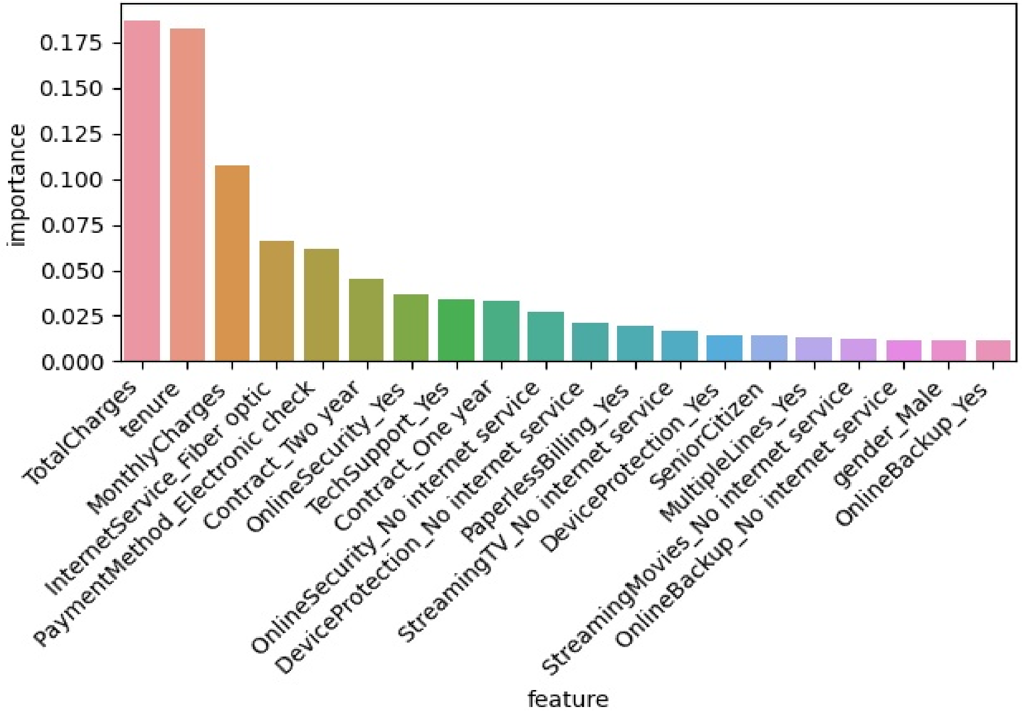
real	No	No	No	No	Yes
prediction	No	No	No	No	Yes
correct	1	1	1	1	1

Incorrectly classified

gender	Male	Male	Male	Female	Male
SeniorCitizen	0.0	0.0	0.0	0.0	0.0
Partner	No	Yes	Yes	No	Yes
Dependents	No	No	Yes	No	Yes
tenure	67.0	17.0	8.0	11.0	19.0
PhoneService	Yes	Yes	Yes	Yes	Yes
MultipleLines	Yes	Yes	No	No	No
InternetService	Fiber optic	Fiber optic	DSL	DSL	Fiber optic
OnlineSecurity	Yes	Yes	No	Yes	No
OnlineBackup	Yes	Yes	No	No	Yes
DeviceProtection	Yes	Yes	Yes	No	No
TechSupport	Yes	No	No	No	Yes
StreamingTV	Yes	No	No	No	Yes
StreamingMovies	Yes	No	No	No	Yes
Contract	Two year	Month-to-month	Month-to-month	Month-to-month	One year
PaperlessBilling	No	No	Yes	No	No
PaymentMethod	Credit card (automatic)	Credit card (automatic)	Electronic check	Credit card (automatic)	Bank transfer (automatic)
MonthlyCharges	116.2	90.2	51.05	48.55	100.0
TotalCharges	7752.3	1454.15	415.05	501.0	1888.65
Churn	Yes	Yes	Yes	Yes	Yes
real	Yes	Yes	Yes	Yes	Yes

prediction	No	No	No	No	No
correct	0	0	0	0	0

Feature importance



The feature importance denotes which features contribute the most to classifying the data. The higher the score, the more relevant is this feature during the decision in which class an entry gets sorted.
Here, we show the 20 most relevant features. In the table below you can find all features.

KOBOLD.AI

Feature	Importance (descending)
TotalCharges	0.1872
tenure	0.1824
MonthlyCharges	0.1076
InternetService_Fiber optic	0.0657
PaymentMethod_Electronic check	0.0619
Contract_Two year	0.0454
OnlineSecurity_Yes	0.037
TechSupport_Yes	0.0341
Contract_One year	0.0329
OnlineSecurity_No internet service	0.0274
DeviceProtection_No internet service	0.0213
PaperlessBilling_Yes	0.0193
StreamingTV_No internet service	0.0165
DeviceProtection_Yes	0.014
SeniorCitizen	0.0139
MultipleLines_Yes	0.0132
StreamingMovies_No internet service	0.0128
OnlineBackup_No internet service	0.0118
gender_Male	0.0116
OnlineBackup_Yes	0.0115

Feature	Importance (descending)
Partner_Yes	0.0098
Dependents_Yes	0.0096
PaymentMethod_Credit card (automatic)	0.0095
StreamingMovies_Yes	0.0088
StreamingTV_Yes	0.0079
InternetService_No	0.0066
TechSupport_No internet service	0.0064
PhoneService_Yes	0.006
PaymentMethod_Mailed check	0.0047
MultipleLines_No phone service	0.0033

Prediction