Analysis of Telecom Churn in Algeria: A Case Study of Ooredoo Using R

BALOULI Houssame Eddine

Abbes Laghrour University Khenchela (Algeria), balouli.houssame.eddine@univ-khenchela.dz

ORCID: https://orcid.org/0009-0006-4015-7450

Received: 06/04/2025 **Accepted: 19/05/2025** **Published:20/07/2025**

Abstract:

This study aims to analyze and predict customer churn in the telecommunications sector, with a particular focus on Ooredoo Algeria. Customer churn—the loss of subscribers—is a critical challenge that directly affects revenue and growth. Using R, we apply machine-learning techniques, specifically Decision Tree and Support Vector Machine (SVM) classifiers, to identify key factors driving churn and build effective models for predicting at-risk customers. Our findings reveal that the Decision Tree model outperforms the SVM model, delivering more accurate and reliable predictions of customer attrition.

Keywords : Customer Churn, Telecommunication, Machine Learning, Decision Tree, Support Vector Machine (SVM).

I- Introduction:

Telecommunications powers the heartbeat of today's economy—fueling innovation, connecting people and businesses, and driving growth across every industry (Kumi et al., 2020; Lam & Shiu, 2010; Paleologos & Polemis, 2013; Symeou, 2011). Telecommunications empowers seamless communication, strengthens businesses, and drives technological innovation that boosts productivity and competitiveness. As a vital pillar of modern infrastructure, the industry not only accelerates the flow of information but also fuels digital transformation across societies—playing a key role in economic growth, job creation, and global progress (Agiakloglou & Gkouvakis, 2015; Giesecke, 2006; Jha & Majumdar, 1999; Lam & Shiu, 2008; Shen, 2012).

Algeria's telecommunications sector is in the midst of a dynamic transformation, shaped by evolving local market trends and rapid global technological progress. The industry is a blend of public and private operators, with the state-owned "Algérie Télécom" continuing to play a dominant role in shaping the market landscape (Yadav, 2016).

The main trend shaping Algeria's telecommunications sector is increased investment in 5G technology, in line with the country's overall drive to digitize the economy as a whole (Hadji, 2023). However, significant challenges remain, most notably the control and integration of artificial intelligence technologies into various services and activities (BENMEHDI & CHOUALI, 2024). While the rollout remains in its early stages—both globally and in Algeria—the move toward 5G promises to dramatically boost connectivity and data service quality. This transition is especially vital as demand for high-speed internet and data continues to surge across the country (Chaabna & Wang, 2015).

Despite its vital role in the economy, the telecommunications sector continues to grapple with the persistent challenge of customer churn—where subscribers discontinue their services. This issue poses a significant threat to telecom providers, affecting revenue, eroding market share, and undermining long-term sustainability. Given that retaining existing customers is often far more cost-effective than acquiring new ones, predicting and managing churn has become a strategic imperative for telecom operators (Dahiya & Bhatia, 2015).

II- Literature review:

A review of churn analysis studies highlights a wide variety of methodologies and datasets, each offering distinct insights. Our study on Ooredoo Algeria employed R, along with Decision Tree and SVM models, applied to a large-scale customer dataset for robust analysis. In contrast, other studies using the same software often relied on smaller datasets, potentially limiting the depth and generalizability of their findings (Lu et al., 2014; Sana et al., 2022) and big data (Ahmad et al., 2019; Xu et al., 2021). Additionally, some researchers opted for Python-based approaches, showcasing the flexibility of tools available for churn analysis across different study contexts (Imani & Arabnia, 2023; Peng et al., 2023; Wassouf et al., 2020). In terms of methodological diversity, some studies limit themselves to a single technique, which may restrict the depth of their analysis and the accuracy of their predictions (Mirkovic et al., 2022). Some other studies adopt a multi-method approach, employing several techniques to compare their effectiveness and ultimately select the best-performing model (Ahmed & Maheswari, 2017; Ehsani & Hosseini, 2024; Jain et al., 2020; Joshi, 2014; Krishna et al., 2024).

III- METHOD

Before we dive into the analysis, it is essential to load the required packages—or install them if they are not already available in the list of installed packages (Fig 1) .For describing our data, we will use the prettyR package, which provides a clear and concise overview of the dataset (Lemon & Grosjean, 2019). Data visualization plays a crucial role in clearly presenting the calculated indicators and various results. To achieve this, we'll utilize the ggplot2 package, which allows for sophisticated and intuitive visualizations (Wickham et al., 2023) and rpart.Plot package (Milborrow, 2024). Since our analysis is based on classification techniques, we will use two key packages. The first is the rpart package, which is essential for building decision tree models (Therneau et al., 2023) and "e1071" package (Meyer [aut et al., 2023). To modify the values of variables and display them as words rather than numbers, we'll use the Catools package, which provides convenient tools for this transformation (Tuszynski, 2021). The remaining functions utilized are basic R functions, which provide essential operations for data manipulation and analysis.

Fig 1: Loaded package

```
library(rpart)
library(rpart.plot)
library(e1071)
library(ggplot2)
library(prettyR)
library(plyr)
library(caTools)
```

Source: R output

The data description (Fig 2) helps us identify the class (type) of each variable, allowing for any necessary adjustments. We use the basic R function "str" to explore our data frame, which is named churn_complete_Data.

Fig 2: Data structure using "str" function

```
str(Churn_Complete_Data)
```

Source: R script

The result is as follow:

Fig 3: The structure of the data

```
$ subscriber_id        : num [1:28188] 52758 101916 102068 119077 138909 ...
$ last_balance         : num [1:28188] 285.13 10.15 10 0.06 272.85 ...
$ rate_plan_description: chr [1:28188] "Haya!" "Haya!" "Haya!" "Haya!" ...
$ Passivity_out        : num [1:28188] 1 0 16 0 40 192 83 0 16 71 ...
$ Passivity_In         : num [1:28188] 0 0 75 0 23 3 84 1 3 4 ...
$ Tenure               : num [1:28188] 187 138 155 182 172 183 183 176 172 181 ...
$ Average_Norm         : num [1:28188] 12.8 2493.6 257.3 2313.4 1105.8 ...
$ Type_HS              : chr [1:28188] "Phablet" "Phablet" "Phablet" "Smartphone" ...
$ Nbre_Achat_M         : num [1:28188] 0 2 0 0 0 0 0 0 0 0 ...
$ CHURN                : num [1:28188] 2 2 2 2 2 2 1 2 2 2 ...
```

Source: R output

The sample size consists of 28,188 observations, representing Ooredoo customers from the state of Khenchela. Our dataset includes eight quantitative (numeric) variables and two nominal (categorical) variables (Fig 3).

The CHURN variable should be a factor rather than numeric. To address this, we made the necessary adjustment. The updated variable is now named Churn_Factor instead of CHURN (Fig 4).

Fig 4: Change CHURN variable from Numeric to Factor

```
Churn_Factor <- as.factor(Churn_Complete_Data$CHURN)
```

Source: R script

We can also assign labels like 'Yes' and 'No' to the new variable instead of '1' and '2'. However, these are just labels—the underlying type of the variable remains unchanged. To achieve this, we use two functions: as.factor and mapvalues (Fig 5).

Fig 5: Add "Yes" and "No" as a wording to the churn factor variable

```
Churn_Complete_Data$CHURN  <-as.factor(mapvalues(Churn_Complete_Data$CHURN, from=c("1","2"), to=c("Yes", "No")))
```

Source: R script

The updated appearance of the CHURN variable is shown below (Fig 6):

Fig 6: The new appearance of the CHURN variable

CHURN
No
No
No
No
No
No
Yes
No

Source: R output

Identifying missing values (Fig 7) and calculating their percentage relative to the total sample size is crucial for ensuring the quality of our results. To do this, we use the following script:

Fig 7: Searching for missing values

```
sapply(Churn_Complete_Data, function(x) sum(is.na(x)))
```

Source: R script

The result is showed in the next figure:

Fig 8: The result of Searching for missing values

```
                      subscriber_id         last_balance rate_plan_description
                                  0                    0                     0
   Passivity_out      Passivity_In               Tenure          Average_Norm
               0                 0                    0                     0
         Type_HS        Nbre_Achat_M               CHURN
               0                 0                    0
```

Source: R output

The results indicate that there are no missing values, allowing us to maintain the full sample size without any adjustments (Fig 8).

Not all variables will be used in our analysis, as some do not contribute meaningful insights. For instance, the subscriber_id variable is excluded from the analysis, which we achieve by assigning it a NULL value (Fig 9).

Fig 9: subscriber_id negligence using NULL function

```
Churn_Complete_Data$subscriber_id <- NULL
```

Source: R script

To enhance the quality of the model, we can split the data into training and test subsets. The following script is used to achieve this (Fig 10):

Fig 10: Divide the sample into two subsets: training and testing

```
set.seed(1)
split_train_test <- sample.split(Churn_Complete_Data$CHURN, SplitRatio = 0.7)
dtrain<- Churn_Complete_Data[split_train_test,]
dtest<- Churn_Complete_Data[-split_train_test,]
```

Source: R script

We begin by describing our variables through tables and graphs, starting with the frequency table for the rate_plan_description variable (Fig 11):

Fig 11: script of the creation of rate_plan_description frequency table

```
freq_table <- Churn_Complete_Data %>%
  group_by(Churn_Complete_Data$rate_plan_description) %>%
  summarise(frequency = n()) %>%
  arrange(desc(frequency))
```

Source: R script

The frequency table of the "rate_plan_description frequency" variable is as follow (Tab 1):

Tab 1: rate_plan_description frequency table

```
> freq_table
# A tibble: 16 x 2
   `Churn_Complete_Data$rate_plan_description` frequency
   <chr>                                           <int>
 1 Haya!                                           14880
 2 La GOLD                                          9357
 3 FREE 200                                          955
 4 Hashta                                            872
 5 La Gold Jdida                                     639
 6 FREE 2000                                         393
 7 Hanya                                             324
 8 LA 1000                                           273
 9 STAR HALA                                         272
10 Nedjma Plus New                                   117
11 Nedjma Free                                        49
12 La Carte Star                                      46
13 La Star 55                                          6
14 Nedjma Bonus                                        2
15 Nedjmax 90                                          2
16 STAR TWEM                                           1
```

Source: R output

From the frequency table (Tab 1) we can see that over 64% of Ooredoo customers are on prepaid plans, which include offerings such as Haya, Free 200, Hashta, Free 2000, Hanya, La 1000, STAR HALA, Nedjma Plus New, Nedjma Free, La Carte Star, La Star 55, Nedjma Bonus, Nedjmax 90, and STAR TWEM. The remaining customers are on postpaid plans, such as La GOLD and La Gold Jdida. This highlights the diversity of Ooredoo's customer base and their preference for the variety of plans available.

The graph provides a clear understanding of the data, especially when dealing with large tables containing numerous variables. Below is the script used to create the bar chart (Fig 12):

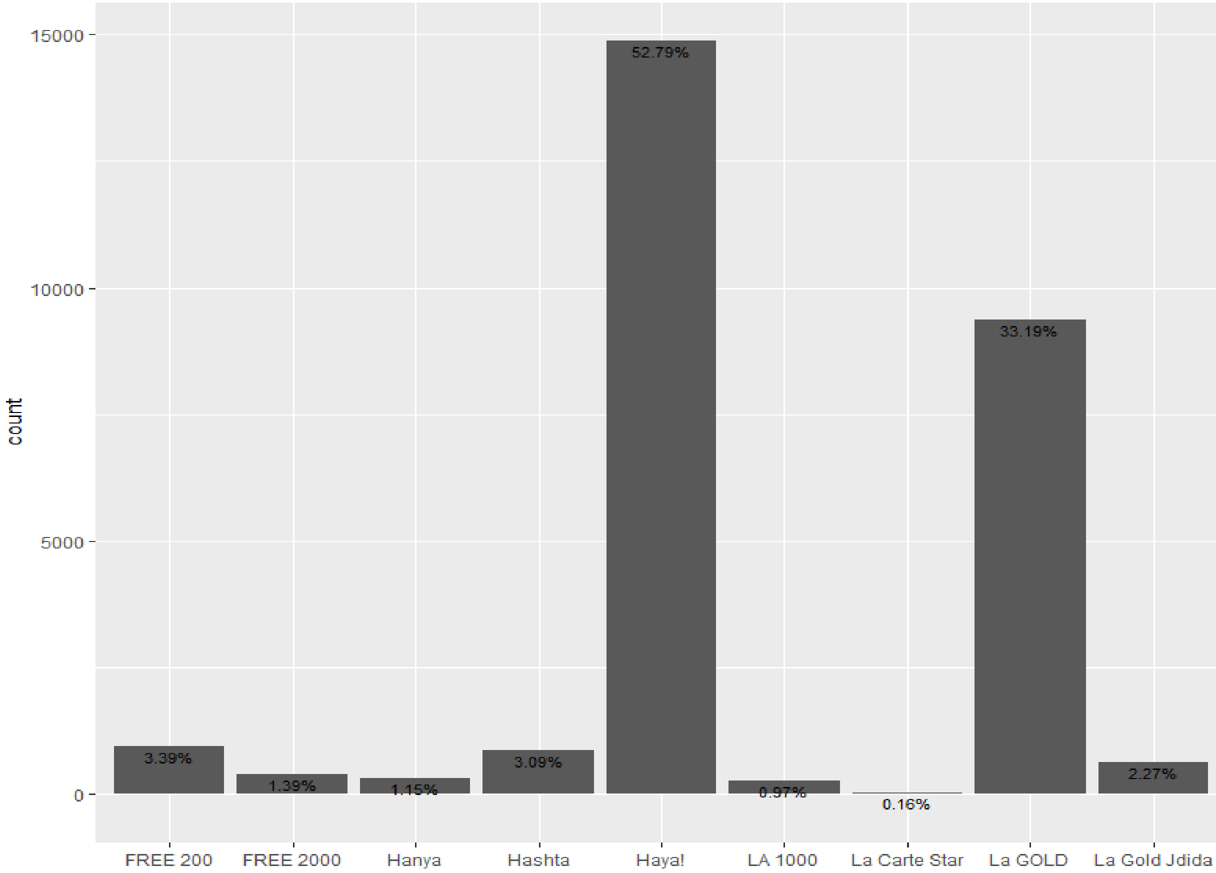
Fig 12: The script of rate_plan_description bar chart

```
p1 <- ggplot(Churn_Complete_Data, aes(x = Churn_Complete_Data$rate_plan_description)) +
  geom_bar() +
  geom_text(aes(y = ..count.. -200,
              label = paste0(round(prop.table(..count..),4) * 100, '%')),
            stat = 'count',
            position = position_dodge(.1),
            size = 3)
p1
```

Source: R script

The result (Fig 13) is as follow with same interpretation as the frequency table but more clearer:

Fig 13: rate_plan_description bar chart



Source: R output

We do the same thing for HS variable (The type of the phone), first the frequency table and then the bar chart (Fig 14):

Fig 14: The script of HS frequency table

```
freq_table_HS <- Churn_Complete_Data %>%
  group_by(Churn_Complete_Data$Type_HS) %>%
  summarise(frequency = n()) %>%
  arrange(desc(frequency))
```

Source: R script

The frequency table of the “HS” variable (Tab 2) is as follow:

Tab 2: HS frequency table

```
> freq_table_HS
# A tibble: 9 x 2
  `Churn_Complete_Data$Type_HS` frequency
  <chr>                             <int>
1 Smartphone                        14177
2 Phablet                            6890
3 Basic Phone                        5151
4 Feature Phone                      1216
5 Tablet                              406
6 Router                              225
7 USB Modem                            74
8 Mobile broadband PCI card            36
9 Dongle                               13
```

Source: R output

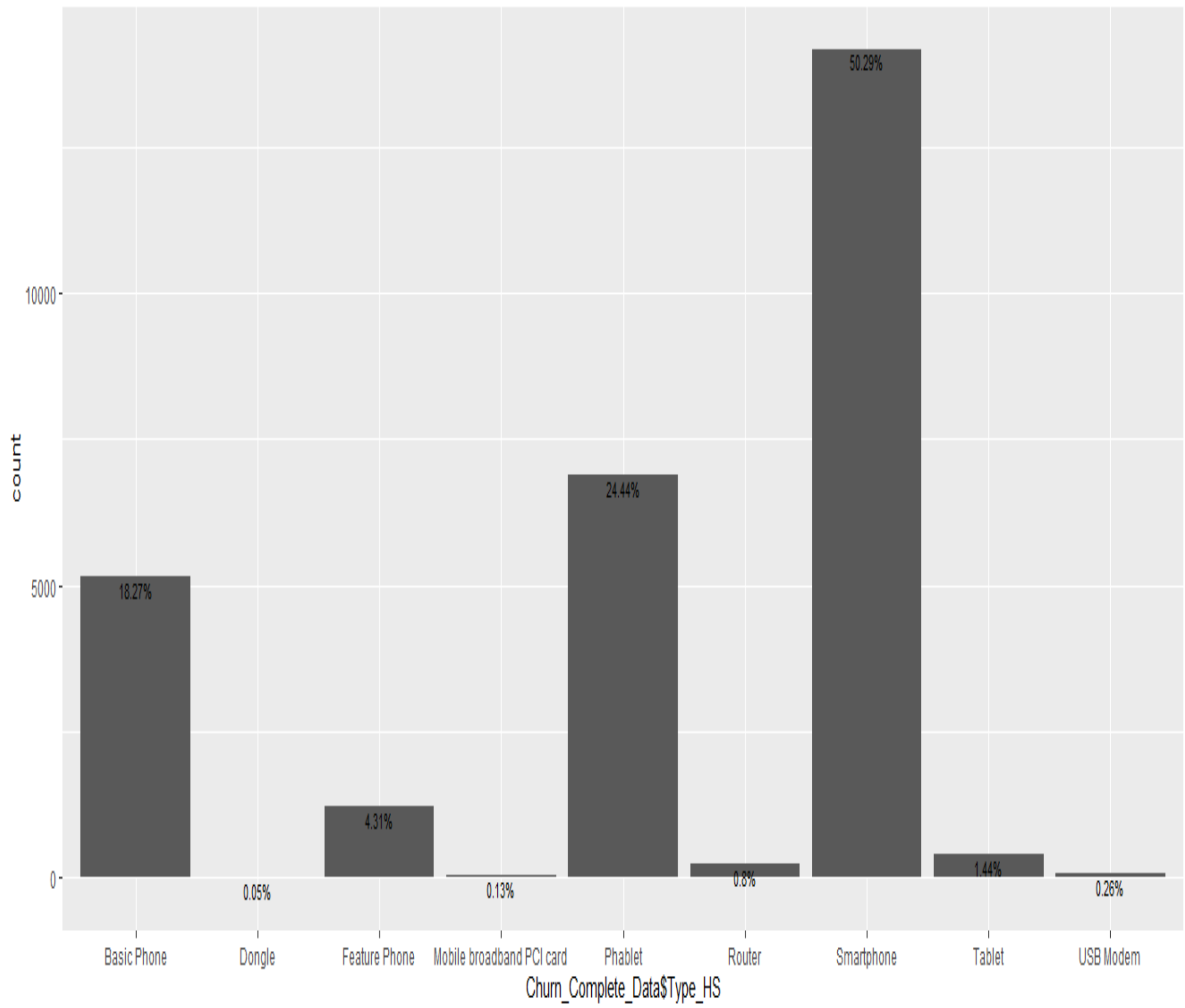
We notice from the frequency table (Tab 2) that Ooredoo customers use 80% of phones that can connect to internet (all types except basic phones). We can see also Fig 16, created by Fig 15.

Fig 15: The script of HS bar chart

```
p2 <- ggplot(Churn_Complete_Data, aes(x = Churn_Complete_Data$Type_HS)) +
  geom_bar() +
  geom_text(aes(y = ..count.. -200,
                label = paste0(round(prop.table(..count..),4) * 100, '%')),
            stat = 'count',
            position = position_dodge(.1),
            size = 3)
```

Source: R script

Fig 16: HS bar chart



Source: R output

The third variable is CHURN.

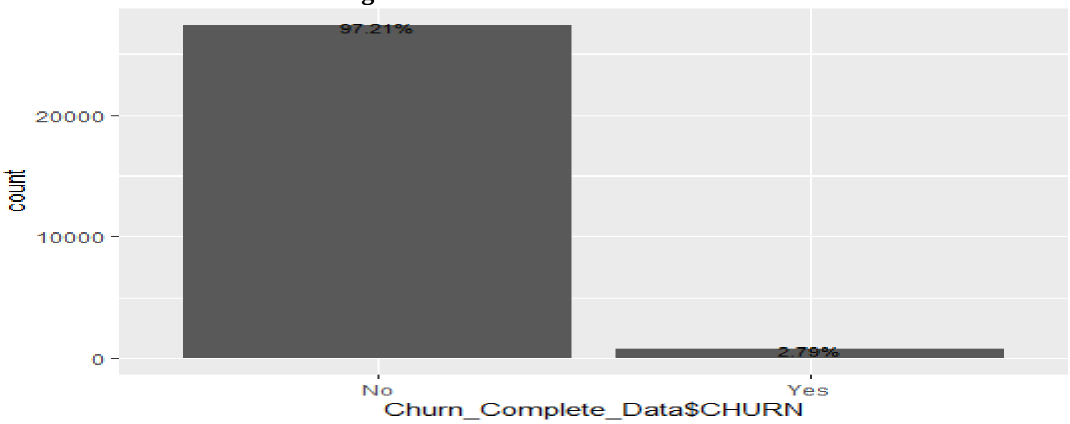
Fig 17: The script of CHURN bar chart

```
p3 <- ggplot(Churn_Complete_Data, aes(x = Churn_Complete_Data$CHURN)) +
  geom_bar() +
  geom_text(aes(y = ..count.. -200,
                label = paste0(round(prop.table(..count..),4) * 100, '%')),
            stat = 'count',
            position = position_dodge(.1),
            size = 3)
```

Source: R script

The bar chart of CHURN variable is as follow:

Fig 18: CHURN bar chart



Source: R output

The percentage of churned customers is very small and acceptable for the analysis. (less than 3%) Fig 18 created by Fig 17.

IV- Results

The first result is the decision tree, using the following script (Fig 19):

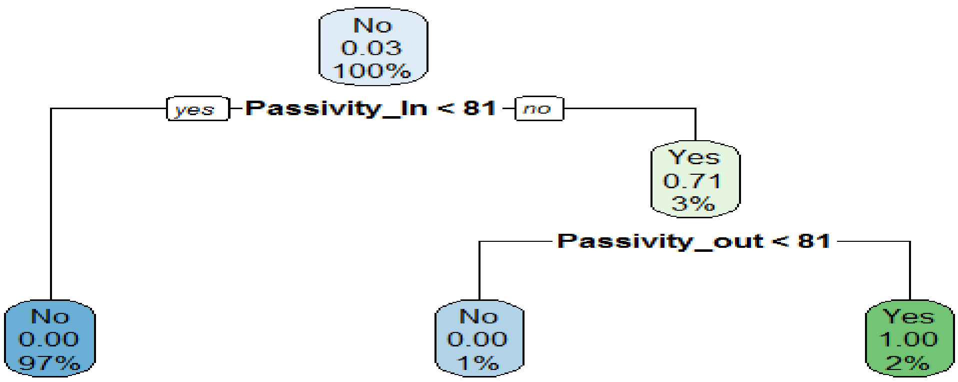
Fig 19: script of the decision tree analysis

```
tr_fit <- rpart(Churn_Complete_Data$CHURN ~., data = dtrain, method="class")
rpart.plot(tr_fit)
```

Source: R script

The result is as follow (Fig 20):

Fig 20: The results of the decision tree analysis



Source: R output

Interpretation :

- **Passivity_In < 81 :**

Most users with incoming passivity (number of passive interactions they receive) less than 81 do not churn. This group constitutes 97% of the data and has a 0.00 probability of churn.

- **Passivity_In >= 81 :**

For users with an incoming passivity of 81 or more:

If their outgoing passivity (the number of passive interactions they make) is less than 81, the probability of churning is 0.71.

If their outgoing passivity is 81 or more, the likelihood of churn is further divided:

Those with lower outgoing passivity (< 81) still have a notable chance of churning.

Those with higher outgoing passivity (> 81) have a 1.00 probability of churning.

The Passivity_In feature is a key indicator of churn risk. Users with high Passivity_In (>= 81) are more likely to churn. Among these users, Passivity_out helps refine the churn prediction: those with lower Passivity_out (< 81) are less likely to churn compared to those with higher Passivity_out

The next element is the confusion matrix (Fig 21), which is used to evaluate the performance of the classification model:

Fig 21: The confusion matrix script

```
tr_prob1 <- predict(tr_fit, dtest)
tr_pred1 <- ifelse(tr_prob1[,2] > 0.5,"Yes","No")
table(Predicted = tr_pred1, Actual = Churn_Complete_Data$CHURN)
```

Source: R script

The result of the confusion matrix is as follow (Fig 22):

Fig 22: The confusion matrix result

```
          Actual
Predicted    No   Yes
      No  27401    88
      Yes     0   699
```

Source: R output

The interpretation of the confusion matrix :

- Actual vs. Predicted :

The rows represent the predicted classes.

The columns represent the actual classes.

- Values in the Matrix :

- True Negatives (TN) : 27,401 (Predicted No, Actual No)
- False Positives (FP): 88 (Predicted No, Actual Yes)
- False Negatives (FN): 0 (Predicted Yes, Actual No)
- True Positives (TP): 699 (Predicted Yes, Actual Yes)

From the confusion matrix, we can calculate several important metrics:

- **Accuracy :**

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Accuracy = (699 + 27,401) / (699 + 27,401 + 88 + 0) = 98.4%

- **Precision** :

Precision = TP / (TP + FP)

Precision = 699 / (699 + 88) = 88.8%

- **Recall (Sensitivity or True Positive Rate) :**

Recall = TP / (TP + FN)

Recall = 699 / (699 + 0) = 100%

- **F1-Score :**

F1-Score = 2 * (Precision * Recall) / (Precision + Recall)

F1-Score = 2 * (0.888 * 1.0) / (0.888 + 1.0) = 94.0%

V- Discussion

- ✓ High Accuracy : The model correctly predicted 98.4% of the instances.
- ✓ High Precision: When the model predicts churn (Yes), it is correct 88.8% of the time.
- ✓ Perfect Recall: The model successfully identified all actual churn cases (100% recall).
- ✓ High F1-Score: A balance between precision and recall, indicating overall good performance.
- ✓ No False Negatives : The model did not miss any actual churn cases.
- ✓ Low False Positives: The model made a small number of errors by predicting non-churned instances as churned.

The model performs exceptionally well, particularly in accurately identifying churned customers (high recall). It also strikes a strong balance with precision, resulting in a high F1-score. This indicates that the model is highly effective at predicting customer churn and can reliably identify at-risk customers with confidence.

The second result is the Support Vector Machine (SVM) classifier

The use of a Support Vector Machine (SVM) classifier (Fig 23) for the churn analysis and its corresponding confusion matrix results.

We start by Training the SVM Classifier:

Fig 23: SVM analysis script

```
classifier = svm(formula = Churn_Complete_Data$CHURN ~ .,
+                         data = dtrain,
+                         type = 'C-classification',
+                         kernel = 'linear')
```

Source: R script

This line trains an SVM classifier using a linear kernel on the training data (dtrain).

The formula "Churn_Complete_Data$CHURN ~ ." specifies that CHURN is the target variable, and all other variables in the dataset are used as predictors.

The next step is making predictions (Fig 24):

Fig 24: Predictions using SVM classifier (script)

```
y_pred = predict(classifier, newdata = dtest)
```

Source: R script

This line uses the trained SVM classifier to make predictions on the test data (dtest).

Creating the Confusion Matrix (Fig 25):

Fig 25: The script of SVM Confusion Matrix

```
cm = table(dtest, y_pred)
```

Source: R script

This line creates a confusion matrix to compare the predicted values (y_pred) with the actual values in the test set (dtest).

Confusion Matrix Fig 26):

Fig 26: SVM Confusion Matrix

```
       y_pred
            1     2
  No    27401     0
  Yes     787     0
```

Source: R output

Interpretation of the SVM Confusion Matrix (Fig 26):

- ✓ True Negatives (TN) : 27,401 (Predicted No, Actual No)
- ✓ False Positives (FP) : 0 (Predicted Yes, Actual No)
- ✓ False Negatives (FN) : 787 (Predicted No, Actual Yes)
- ✓ True Positives (TP) : 0 (Predicted Yes, Actual Yes)

Performance Metrics Calculation :

- ➢ **Accuracy :**

Accuracy = 97.22%

- ➢ **Precision** :

Precision = 0

- ➢ **Recall (Sensitivity or True Positive Rate) :**

Recall = 0%

- ➢ **F1-Score :**

F1-Score = 0

VI- Disscussion

High Accuracy: The model achieved an impressive accuracy of 97.22%, largely due to its correct predictions of non-churn cases.

Zero Precision and Recall for the Churn Class: However, the model fails to correctly predict any of the actual churn cases, resulting in zero precision and recall for the 'Yes' (churn) class.

Misclassification of Churn Cases: All 787 actual churn instances are misclassified as non-churn, highlighting the model's ineffectiveness in identifying churned customers.

Despite the high overall accuracy, the SVM classifier with a linear kernel has failed to identify any churned customers. This poor performance in predicting the churn class suggests that the model is not suitable for churn prediction in this context, and alternative approaches or further model tuning may be required.

The Decision Tree model (Fig 22) significantly outperforms the SVM model (Fig 26) in predicting customer churn. It provides high accuracy, precision, recall, and F1-score, making it a more reliable choice for this analysis. The SVM model, in its current form, is not effective in identifying churned customers, and would require substantial improvements or reconsideration of the modeling approach.

VII- Conclusion:

The telecommunication sector, a cornerstone of modern economies, faces considerable challenges, with customer churn being one of the most impactful, as it directly affects both revenue and growth. This study aimed to analyze and predict customer churn at Ooredoo Algeria using machine learning techniques implemented in R. The process included data preparation, analysis through Decision Tree and Support Vector Machine (SVM) classifiers, and subsequent model evaluation. The Decision Tree model outperformed the SVM model, showing significantly higher accuracy and reliability.

Key insights from the analysis underscored the importance of customer behaviors, particularly incoming and outgoing passivity, in predicting churn. The Decision Tree model achieved an impressive 98.4% accuracy, with perfect recall, demonstrating its capability in identifying at-risk customers. In contrast, while the SVM model displayed high overall accuracy, it failed to correctly predict churn, proving less effective in this scenario.

Based on these findings, it is recommended that Ooredoo Algeria prioritize monitoring customer passivity as a key churn indicator. Additionally, implementing targeted retention strategies for customers identified as high-risk by the Decision Tree model could significantly mitigate churn rates. Further research could involve integrating more advanced machine learning techniques and larger datasets to improve the predictive power and robustness of the models

Bibliography

Agiakloglou, C., & Gkouvakis, M. (2015). Causal interrelations among market fundamentals : Evidence from the European Telecommunications sector. The Quarterly Review of Economics and Finance, 55, 150‑159. https://doi.org/10.1016/j.qref.2014.07.001

Ahmad, M., Tang, X.-W., Qiu, J.-N., & Ahmad, F. (2019). Interpretive Structural Modeling and MICMAC Analysis for Identifying and Benchmarking Significant Factors of Seismic Soil Liquefaction. Applied Sciences, 9(2), 233. https://doi.org/10.3390/app9020233

Ahmed, A. A. Q., & Maheswari, D. (2017). Churn prediction on huge telecom data using hybrid firefly based classification. Egyptian Informatics Journal, 18(3), 215‑220. https://doi.org/10.1016/j.eij.2017.02.002

BENMEHDI, S., & CHOUALI, A. (2024). L'intelligence artificielle IA transforme le monde des affaires. مجلة الدراسات التجارية والاقتصادية المعاصرة, 7(1), 245‑257.

Chaabna, S., & Wang, H. (2015). Analysis of the State of E-commerce in Algeria. International Journal of Marketing Studies, 7(2), p44. https://doi.org/10.5539/ijms.v7n2p44

Dahiya, K., & Bhatia, S. (2015). Customer churn analysis in telecom industry. 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), 1‑6. https://doi.org/10.1109/ICRITO.2015.7359318

Ehsani, F., & Hosseini, M. (2024). Customer churn analysis using feature optimization methods and tree-based classifiers. Journal of Services Marketing, 39(1), 20‑35. https://doi.org/10.1108/JSM-04-2024-0156

Giesecke, J. (2006). The National and Regional Economic Consequences of Rapid Growth in Australia's Telecommunications Sector. Economic Analysis and Policy, 36(1), 61‑97. https://doi.org/10.1016/S0313-5926(06)50005-0

Hadji, S. (2023). The digital economy in Algeria—Prospects and challenges. مجلة الدراسات التجارية والاقتصادية المعاصرة, 6(2), 174‑187.

Imani, M., & Arabnia, H. R. (2023). Hyperparameter Optimization and Combined Data Sampling Techniques in Machine Learning for Customer Churn Prediction : A Comparative Analysis. Technologies, 11(6), 167. https://doi.org/10.3390/technologies11060167

Jain, H., Khunteta, A., & Srivastava, S. (2020). Churn Prediction in Telecommunication using Logistic Regression and Logit Boost. Procedia Computer Science, 167, 101‑112. https://doi.org/10.1016/j.procs.2020.03.187

Jha, R., & Majumdar, S. K. (1999). A matter of connections : OECD telecommunications sector productivity and the role of cellular technology diffusion. Information Economics and Policy, 11(3), 243‑269. https://doi.org/10.1016/S0167-6245(99)00017-7

Joshi, S. (2014). Customer Experience Management : An Exploratory Study on the Parameters Affecting Customer Experience for Cellular Mobile Services of a Telecom Company. Procedia - Social and Behavioral Sciences, 133, 392‑399. https://doi.org/10.1016/j.sbspro.2014.04.206

Krishna, R., Jayanthi, D., Shylu Sam, D. S., Kavitha, K., Maurya, N. K., & Benil, T. (2024). Application of machine learning techniques for churn prediction in the telecom business. Results in Engineering, 24, 103165. https://doi.org/10.1016/j.rineng.2024.103165

Kumi, E., Yeboah, T., & Kumi, Y. A. (2020). Private sector participation in advancing the Sustainable Development Goals (SDGs) in Ghana : Experiences from the mining and telecommunications sectors. The Extractive Industries and Society, 7(1), 181‑190. https://doi.org/10.1016/j.exis.2019.12.008

Lam, P.-L., & Shiu, A. (2008). Productivity analysis of the telecommunications sector in China. Telecommunications Policy, 32(8), 559‑571. https://doi.org/10.1016/j.telpol.2008.06.004

Lam, P.-L., & Shiu, A. (2010). Economic growth, telecommunications development and productivity growth of the telecommunications sector : Evidence around the world. Telecommunications Policy, 34(4), 185‑199. https://doi.org/10.1016/j.telpol.2009.12.001

Lemon, J., & Grosjean, P. (2019). prettyR : Pretty Descriptive Stats (Version 2.2-3) [Logiciel]. https://cran.r-project.org/web/packages/prettyR/index.html

Lu, N., Lin, H., Lu, J., & Zhang, G. (2014). A Customer Churn Prediction Model in Telecom Industry Using Boosting. IEEE Transactions on Industrial Informatics, 10(2), 1659‑1665. https://doi.org/10.1109/TII.2012.2224355

Meyer [aut, D., cre, Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., C++-code), C.-C. C. (libsvm, & C++-code), C.-C. L. (libsvm. (2023). e1071 : Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien (Version 1.7-14) [Logiciel]. https://cran.r-project.org/web/packages/e1071/index.html

Milborrow, S. (2024). rpart.plot : Plot « rpart » Models: An Enhanced Version of « plot.rpart » (Version 3.1.2) [Logiciel]. https://cran.r-project.org/web/packages/rpart.plot/index.html

Mirkovic, M., Lolic, T., Stefanovic, D., Anderla, A., & Gracanin, D. (2022). Customer Churn Prediction in B2B Non-Contractual Business Settings Using Invoice Data. Applied Sciences, 12(10), 5001. https://doi.org/10.3390/app12105001

Paleologos, J. M., & Polemis, M. L. (2013). What drives investment in the telecommunications sector? Some lessons from the OECD countries. Economic Modelling, 31, 49‑57. https://doi.org/10.1016/j.econmod.2012.11.038

Peng, K., Peng, Y., & Li, W. (2023). Research on customer churn prediction and model interpretability analysis. PLOS ONE, 18(12), e0289724. https://doi.org/10.1371/journal.pone.0289724

Sana, J. K., Abedin, M. Z., Rahman, M. S., & Rahman, M. S. (2022). A novel customer churn prediction model for the telecommunication industry using data transformation methods and feature selection. PLOS ONE, 17(12), e0278095. https://doi.org/10.1371/journal.pone.0278095

Shen, W. (2012). Deconstructing the myth of Alipay Drama—Repoliticizing foreign investment in the telecommunications sector in China. Telecommunications Policy, 36(10), 929‑942. https://doi.org/10.1016/j.telpol.2012.08.008

Symeou, P. C. (2011). Economy size and performance : An efficiency analysis in the telecommunications sector. Telecommunications Policy, 35(5), 426‑440. https://doi.org/10.1016/j.telpol.2011.03.008

Therneau, T., Atkinson, B., port, B. R. (producer of the initial R., & maintainer 1999-2017). (2023). rpart : Recursive Partitioning and Regression Trees (Version 4.1.23) [Logiciel]. https://cran.r-project.org/web/packages/rpart/index.html

Tuszynski, J. (2021). caTools : Tools: Moving Window Statistics, GIF, Base64, ROC AUC, etc (Version 1.18.2) [Logiciel]. https://cran.r-project.org/web/packages/caTools/index.html

Wassouf, W. N., Alkhatib, R., Salloum, K., & Balloul, S. (2020). Predictive analytics using big data for increased customer loyalty : Syriatel Telecom Company case study. Journal of Big Data, 7(1), 29. https://doi.org/10.1186/s40537-020-00290-0

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., Posit, & PBC. (2023). ggplot2 : Create Elegant Data Visualisations Using the Grammar of Graphics (Version 3.4.4) [Logiciel]. https://cran.r-project.org/web/packages/ggplot2/index.html

Xu, T., Ma, Y., & Kim, K. (2021). Telecom Churn Prediction System Based on Ensemble Learning Using Feature Grouping. Applied Sciences, 11(11), 4742. https://doi.org/10.3390/app11114742

Yadav, N. (2016). Analysis of telecom industry in Algeria. http://repository.iimb.ac.in/handle/123456789/12532