Pinky: churn rate risk

The client

"Pinky" is a telecommunications company selling mobile and internet packages. The clients are monthly subscribers but some with annual or biannual contracts.

The Problem

Every month, a lot of clients quit their subscriptions and add up to the "churn rate" of "Pinky Telecom". To reduce this rate, the company requested help and made available its clients data from a given month and demanded an analysis on the clients with the goal of anticipating the ones who might be next to quit and to try to avoid it.

"Pinky Telecom" has a call center with commercial assistants that can call the clients and offer them new proposals. The company noticed that it is almost impossible to recover clients who have canceled their subscriptions and expects to convince the clients "at risk" to renew before they quit the services.

The RESPONSE and recommendation:

Having that in mind, after our analysis, we predicted about 30 active "Pinky" clients with more than 60% of chances to "churn" and with whom "Pinky" commercial team might consider an approach to avoid them leaving.

These clients are active for more than 4 months and for 13 months on average.
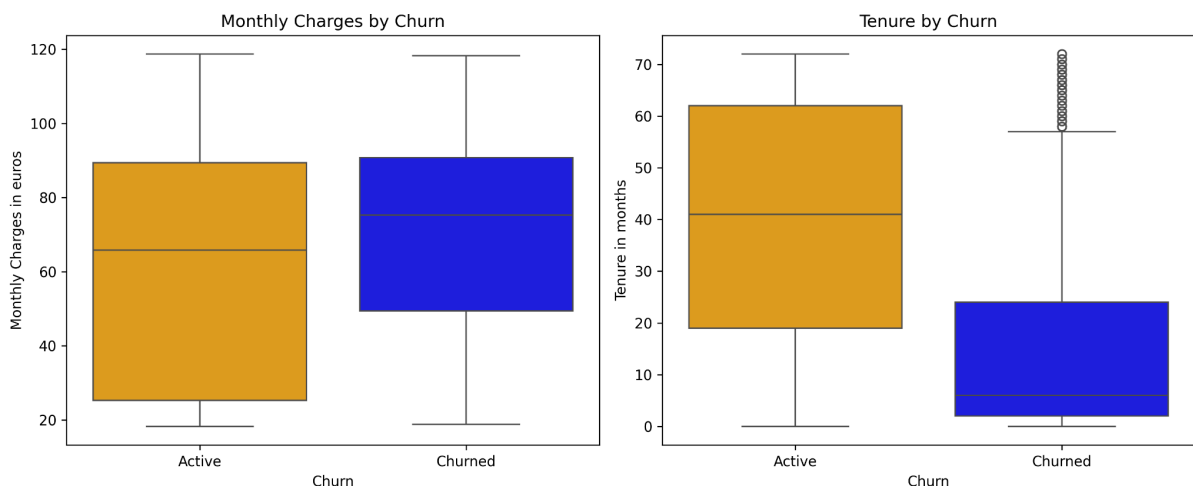
The majority is linked to an Internet DSL service, some have also the Phone Service. The average of Monthly Charges of this group is around 55 euros and at least one who has been with the company for almost 4 years (46 months) is paying nearly 44 euros a month, for DSL with streamingTV and streamingMovies, showing 60% of chances of "churning".

At least 9 of our 30 high risk clients have declared to have a partner and 5 have dependents, including one without a partner, which shows that family/collective packages might affect this clients decisions to subscribe to services.

From our analysis we can underline that **a lot of the churn rate is linked to "Month-to-Month"** type of contracts and **internet clients without extra services** which might have helped to convince others to stay with Pinky services.
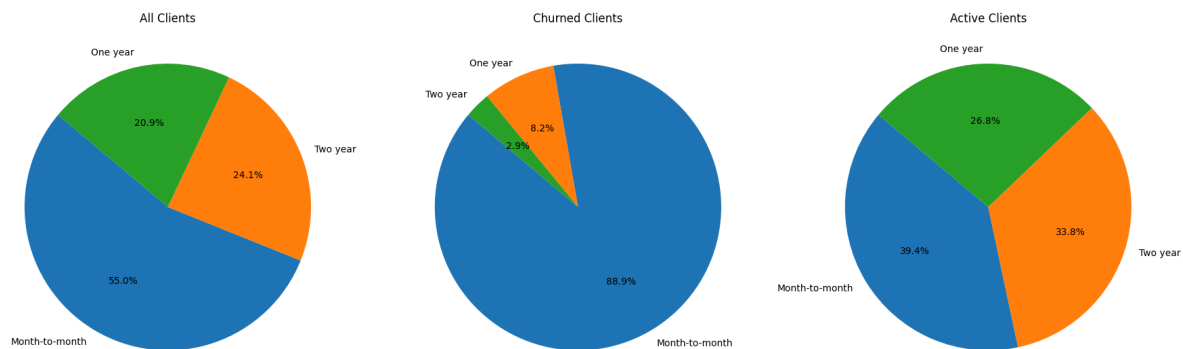
The majority of the **31% of churn clients apparently take decisions alone** and don't have dependents. At least **75% of them quit before they complete 2 years** with Pinky services which might be linked to the end contracts and easy to quit "month-to-month" services.

The personalization of packages and/or the renegotiation of services can be a way to retain clients that might be always looking for cheaper conditions in a competitive market.

Today, almost no one wants to be without mobile or internet but if they can save some money, the bureaucracy of changing services might be less important than the savings at the end of the month.

The dataset made available

We got a dataset with the info of 7043 not directly identified clients because of the General Data Personal Regulation (GDPR). Either way we had access to some personal info like gender; seniority, civil status and if they had dependents which allowed us to have an idea if family could be on the decision table for the kind of services we are working for.

From the data, we noted that there were a bit more male clients (3555) than women (3488), the large majority was under 65 years old, didn't had any dependents (4933) but it was balanced between the ones with or without a declared partner (3402/3641).

The Exploratory Data Analysis

We started by checking the data made available, looking for some odd values and how many numerical ones we had, noticing the 7043 unique entries in the dataset.

From the 21 features, including the anonymous "customerID" code, only 3 were numerical. A column with the total of all the money already paid by the client to Pinky was not numerical but after observing we decided it was not pertinent for our analysis.

There were no empty values. The clients were balanced between women and men although the testosterone was slightly higher. Clients with more than 65 years old (seniors) were less than 17%, 560 were alone, 570 had a partner and 90 had dependents.

At least 4824 of the clients were still active with subscriptions and 2219 had "churned", at a rate of 31,5%.

Going into data visualization (DataViz), we noticed the close correlations between "tenure" and "monthly" contracts with the churn rate. We can not point to a cause-effect link but they're connected.

The load of the monthly charge might have a heavier load in the decision for quitting as we could notice that at least 50% of the churn rate were made by clients that were paying between 45 and 90 euros per month, the large majority were DSL clients.

Gender, on the other hand, had no role in the churn rate as it is balanced both ways.

Family seems to play a role as clients with partners and or dependents seem less inclined to churn.

The phone service is well present in subscriptions but its DSL the service with more hits in the churn rate.On the other way, internet fiber clients are more likely to stay, the data shows.

Machine Learning

With all the data analysed we prepared it to be managed by some models of Machine Learning with the goal to predict the active clients in higher risk of "churning".

We had to convert the categorical features into numeric ones and we started by training a model of Logistic Regression as we intended to classify clients by churn risk.

The first results were not bad, but the Recall for "churn" was at 70%. As we needed some improvement in predicting clients who might churn we went on to train a Random Forest, which is known to trust on independent trees to calculate eventual nonlinear correlations.

The correlation matrix showed improvement on the RF predictions.

We applied a GridSearch on the "Forest" to check where we could improve the model. After optimizing the trees and also adding a strong feature like the "monthly" contracts that had

been taken out when we applied the "One-Hot Encoding" Dummies to the "contracts" column, the accuracy overall improved to 88%.

The Recall sensitivity reached 76%.

After this we did another small data analysis to compare active clients and churned, noticing that 23% of the "churned" were senior clients and 55% of the clients with partners and 35% with dependents were more inclined to a stable subscription.

Almost 90% of the "churn" but also of the active clients had the phone service which indicates that DSL internet service is the heavyweight in the "churn" team.

The prediction of our best model identified a **group of 30 active clients showing between 60% to 86% of chances of quitting** Pinky services. A recommendation was issued to approach these clients and avoid that decision to be taken.