Mobile operators sell handsets in large quantities. To be competitive versus the open market and the competitors it is possible to buy new subscription and device by paying the price of device in monthly installments. In addition in case of loyalty contracting for price plan with higher content handset selling price decreases while monthly fee increases by handset premium amount. Due to low level of starting price payed for device there is a share of fraudster contracts and another share of contracts made by customers who will be not able to pay the higher monthly fees and installments and will churn with nopay reason.

Using the attached set of sales events and attributes your tasks are the following:

1. Based on the described situation define a set of business tasks, goals and a set of hypothesis
2. Create exploratory analysis of attached dataset
3. Create predictive model or models which can detect fraudelent clients in the moment of purchase.
   1. Please, give a detailed description of what you did, and why you did. We are curious about how do you think and how do you break down the problem.
   2. Evaluate the model(s), give a short description about the performance of your models.
4. Please, provide a document detailing all non-obvious steps and your answers. You should also provide a script which contains the work that you made on the dataset.

The dataset contains sales events, possible predictors for the outcomes, and target variables (fraud\_status\_6months,nopay\_after\_12months). You can find the variable descriptions in the following table:

|  |  |
| --- | --- |
| **Variable name** | **Variable description** |
| day\_id | Day of purchase |
| id | Purchase ID |
| pl\_subseg\_desc | Client's segment |
| gender | Client's gender |
| r\_age\_y | Client's age |
| instalment\_ind | Indicator of installment |
| instal\_cnt | Length of installment period in months |
| product\_name | Name of the product |
| manufacturer\_name\_en | Product manufacturer |
| operating\_system | Operating system used by product |
| handset\_feature\_cat\_desc | Product capability category |
| moving\_average\_price\_amt | Purchase price of the handset |
| selling\_price\_amt | Selling price of the handset |
| upfront\_pym\_amt | Upfront payment (Paid by client at the purchase) |
| monthly\_fee | Montly fee of the tariff |
| lmh\_desc | Category of the tariff |
| channel\_class | Type of sales channel |
| channel\_group | Type of store |
| address\_county | County of the client |
| outlet\_county | County of the outlet where the purchase happened |
| fraud\_status\_6month | Status of the client after 6 months (Y: fraudelent client, N: active client) |
| nopay\_after\_12month | Status of the client after 12 months (Y: nopay client, N: active client) |