Predicting the probability of user churn for a network of fitness centers

Project status: *completed*

Libraries used: `pandas, matplotlib, seaborn, numpy, datetime, sys, sklearn, scipy, itertools`

Description of the project:

The network of fitness centers "Bodybuilder-datascientist" develops a strategy for interacting with clients based on analytical data.

A common problem for fitness clubs and other services is the outflow of customers. For a fitness center, we can consider that a client has been churned if he has never visited the gym in the last month. Of course, special cases are not excluded, after the elimination of which the client will definitely continue to go to fitness. However, the opposite is more often the case. If a client went to the gym a little, and then disappeared, most likely he will not return.

To combat churn, Bodybuilder Data Scientist's customer service department has digitized many client profiles. Data Bodybuilder provided the information in a csv file that contains the data for the month before the churn and the fact of the churn for a particular month.

Our task is to analyze and prepare an action plan for customer retention.

Research progress:

1. Data review and data preprocessing.
2. Churn probability forecast (at the level of the next month) for each client.
3. Formation of typical client portraits:
    o highlight some of the most prominent groups;
    o to characterize their main properties.
4. Analysis of the main features that most strongly affect the outflow.
5. Formulation of conclusions and development of recommendations for improving the quality of work with clients:
    o identify target customer groups;
    o propose measures to reduce churn;
    o to define other features of interaction with clients.

Conclusions on the project:

The groups of clients in terms of the share of the churn of each group of the total number of churned clients differ quite strongly:

- the strongest outflow occurs in the third group - more than 51%;
- slightly more than 44% in the second group;
- group No. 1 in third place in terms of outflow - almost 27%;
- in the fourth and fifth outflow is not so strong - ~ 7% and ~ 3% respectively (you can consider clients in these groups more reliable).

In the first group clients for whom the club is located close enough, they are often employees of the club's partners or came through a promotion, they have a subscription, on average, for more than half a year, they attend group classes more often than others and spend quite a lot on related services of the club. Almost 5 months have passed since the first application, and subscriptions are still far away, on average, about 9 months. They visit the club around twice a week.

In the second group clients for whom the club is also close enough, most of them have a subscription for up to 6 months, in terms of service consumption, they are somewhere in the middle between all groups, until the end of the subscription, on average, almost 5 months, but from the moment of application almost 4 months have passed, less than half of them attend group clubs, are employees of the club's partners or came through a promotion, visit the gym about 2 times a week.

In the third group clients who are far from the club, less than half of them are employees of the club's partners, monthly subscriptions are more common than for longer periods, they are less likely to go to group classes, but related services consume, go once a week rather than more.

In the fourth group the clients who work/live closest to the club, there are not so many promotional or partner ones among them, subscriptions are preferred for short-term and services are used the least, groups visit, but not often, and they go to the club mostly once a week.

In the fifth group, clients who are also based near the club are not very often partner or promotional clients, but among them more often there are lovers of longer subscriptions or attending group ones, they also have a higher indicator of time from the moment of the first application and the indicator consumption of related services of the club, besides, they are the most frequent clients of the club by visits - almost 3 times a week.