**Student:** Samgat Baknur 20b030663

**Topic:** predicting of car prices in the Kazakhstani market.

A car price prediction model can be very useful in several areas. These are insurance for car valuation, banks, pawnshops and microfinance organizations for valuation of collateral, etc. That is, we can say that there is a demand for this model and it will be relevant.

**Dataset**:

The dataset was parsed by me from the site koles.kz, about a year ago for one freelance order, there are 179063 unique ads from the site in the dataset.

Target variable is price in tenge.

I also added a variable as Score which is the sum of all binary variables, which in my opinion should show good results.

**Challenge**:

I used raw and unprepared data that I collected myself, so I had to clean and prepare the data myself during the development for training the model.

**Algorithm**:

I decided to use the Light gbm library - Gradient Boosting Decision Tree algorithm, as it has a large number of advantages, among which: it does not care about high correlation between variables, it can accept a large number of variables, it can work with categorial variables, but the biggest advantage is its speed.

**Evaluation method:**

I used mape metric. MAPE (Mean Absolute Percentage Error) is a commonly used performance metric in machine learning that measures the accuracy of a model's predictions. It is calculated as the average percentage difference between the predicted and actual values. The formula for MAPE is:

MAPE = (1/n) \* Σ(|actual - predicted| / |actual|) \* 100

where n is the number of samples in the dataset.

**References**: Light GBM documentation - <https://lightgbm.readthedocs.io/en/latest/index.html>