Моделирование осуществляется при помощи методов машинного обучения. Для исследования проблемы нами выбраны 18 моделей в том числе:

* Linear Regression
* Logistic Regression
* Perceptron
* Linear SVC
* MLPClassifier
* Decision Tree Classifier 1
* Stochastic Gradient Decent
* RidgeClassifier
* BaggingClassifier
* AdaBoostClassifier 1
* GradientBoostingClassifie
* KNeighborsClassifier
* DecisionTreeClassifier 2
* RandomForestClassifier
* XGBClassifier
* AdaBoostClassifier 2
* Naive Bayes
* SVC

В результате работы каждой модели выдаются следующая информация в форме таблиц и графиков:

* Таблица классификация
* Матрица Confusion Matrix
* ROC график
* Score график

Таблица классификация даёт оценку модели по критериям:

* ***precision***
* ***recall***
* ***f1-score***

ROC график дапёт оценку модели по критерию AUC, т.е. значению площади под кривой. Что означает – чем ближе этот показатель к значению равному 1, тем выше сответствие прогнозных значений модели реальным входным данным.

Исследовательский анализ включает в себя изучение данных и поиск связей между переменными, которые ранее были неизвестны. Вот что вам нужно знать