

Московский авиационный институт
(национальный исследовательский университет)

Факультет информационных технологий и прикладной математики

Кафедра вычислительной математики и программирования

Лабораторная работа №1 по курсу
«Искусственный интеллект»

Студент: Я. А. Графчикова
Группа: М8О-408Б

Москва, 2019

Постановка задачи

Познакомиться с платформой Azure Machine Learning, реализовав полный цикл разработки решения задачи машинного обучения, используя три различных алгоритма, реализованные на этой платформе.

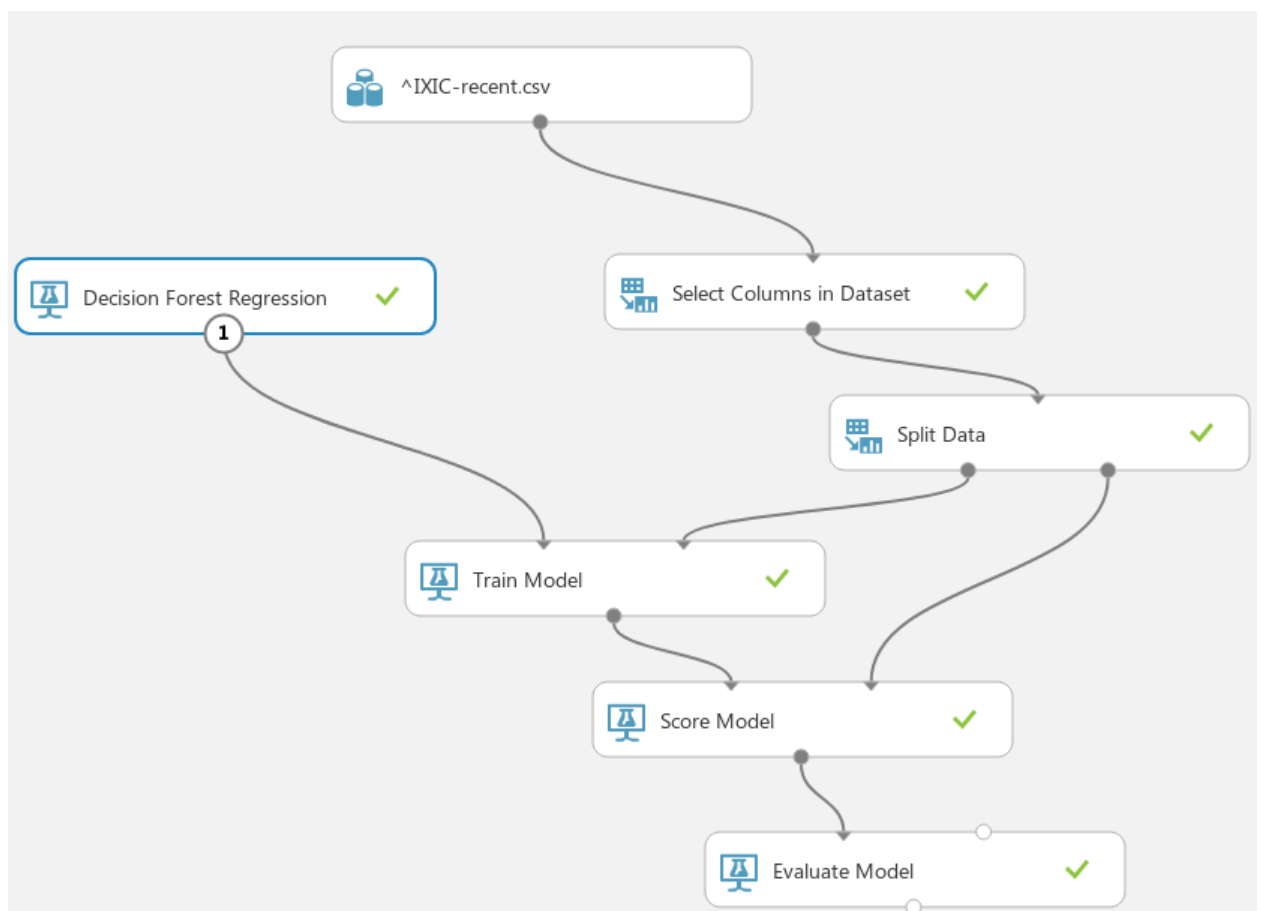
Описание

В данной лабораторной работе я работала с датасетом – история акций компании NASDAQ Composite.

Было использовано 3 алгоритма – Decision Forest Regression, Linear Regression, Two-class Logistic Regression.

Decision Forest Regression

Оценка модели леса решений для задачи регрессии.



Результаты:

Акции алгоритм 2 > Evaluate Model > Evaluation results

rows 1
columns 6

| | Negative Log Likelihood | Mean Absolute Error | Root Mean Squared Error | Relative Absolute Error | Relative Squared Error | Coefficient of Determination |
|---------|-------------------------|---------------------|-------------------------|-------------------------|------------------------|------------------------------|
| view as | 4322.93689 | 15.253956 | 23.302455 | 0.011599 | 0.000201 | 0.999799 |

Linear Regression

Линейная регрессия исследует зависимость одной переменной от нескольких других переменных с линейной функцией зависимости. Для этого в данной модели используется метод наименьших квадратов.

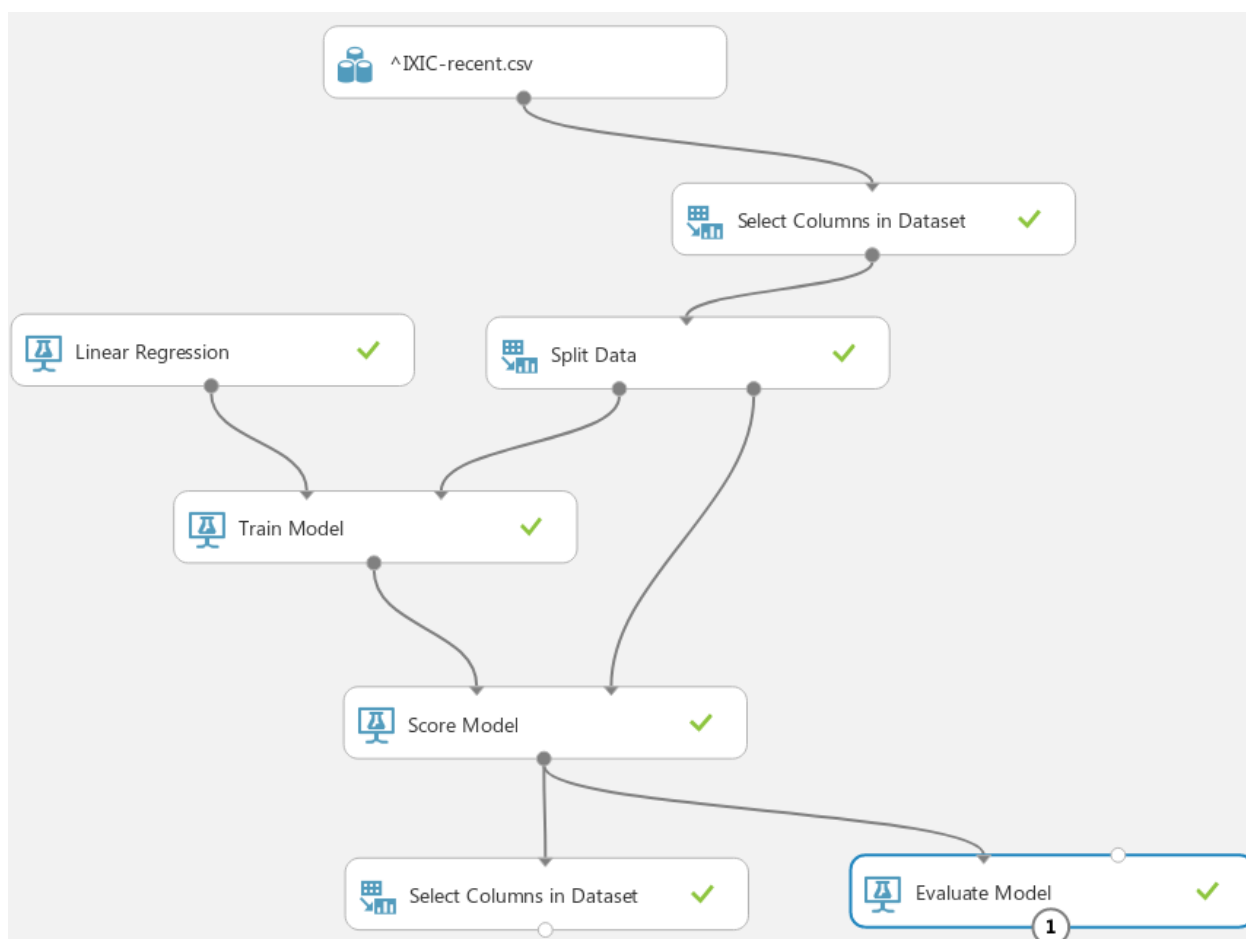
Linear Regression

Solution method

Ordinary Least Squares

L2 regularization weight

0.001



Результаты:

Акции алгоритм 1 ➤ Evaluate Model ➤ Evaluation resu

Metrics

| | |
|------------------------------|-----------|
| Mean Absolute Error | 9.71516 |
| Root Mean Squared Error | 14.357516 |
| Relative Absolute Error | 0.00717 |
| Relative Squared Error | 0.000073 |
| Coefficient of Determination | 0.999927 |

Error Histogram

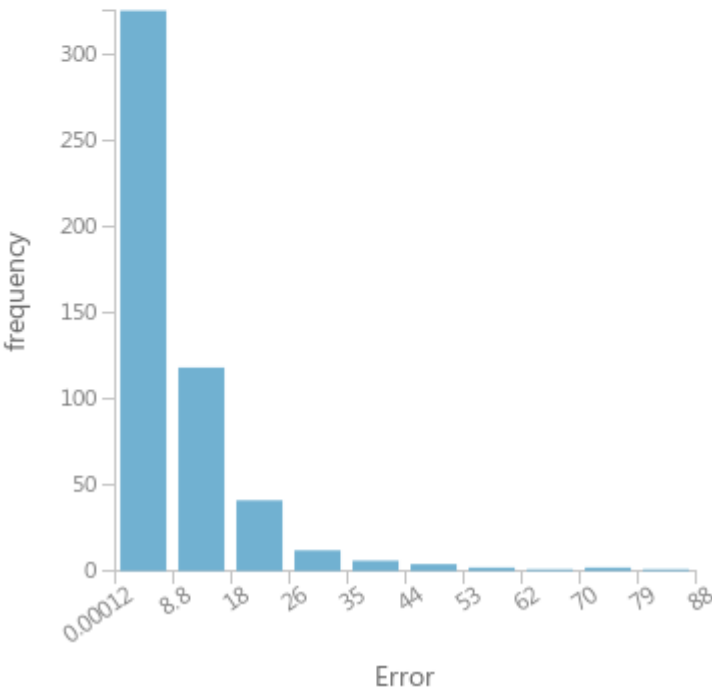


График зависимости полученных значений от ожидаемых:

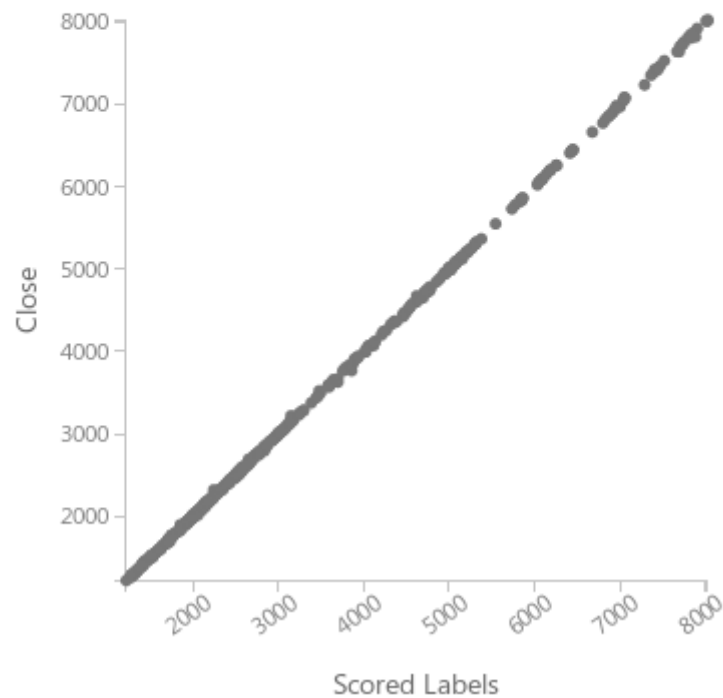
Visualizations

Scored Labels

ScatterPlot

compare to

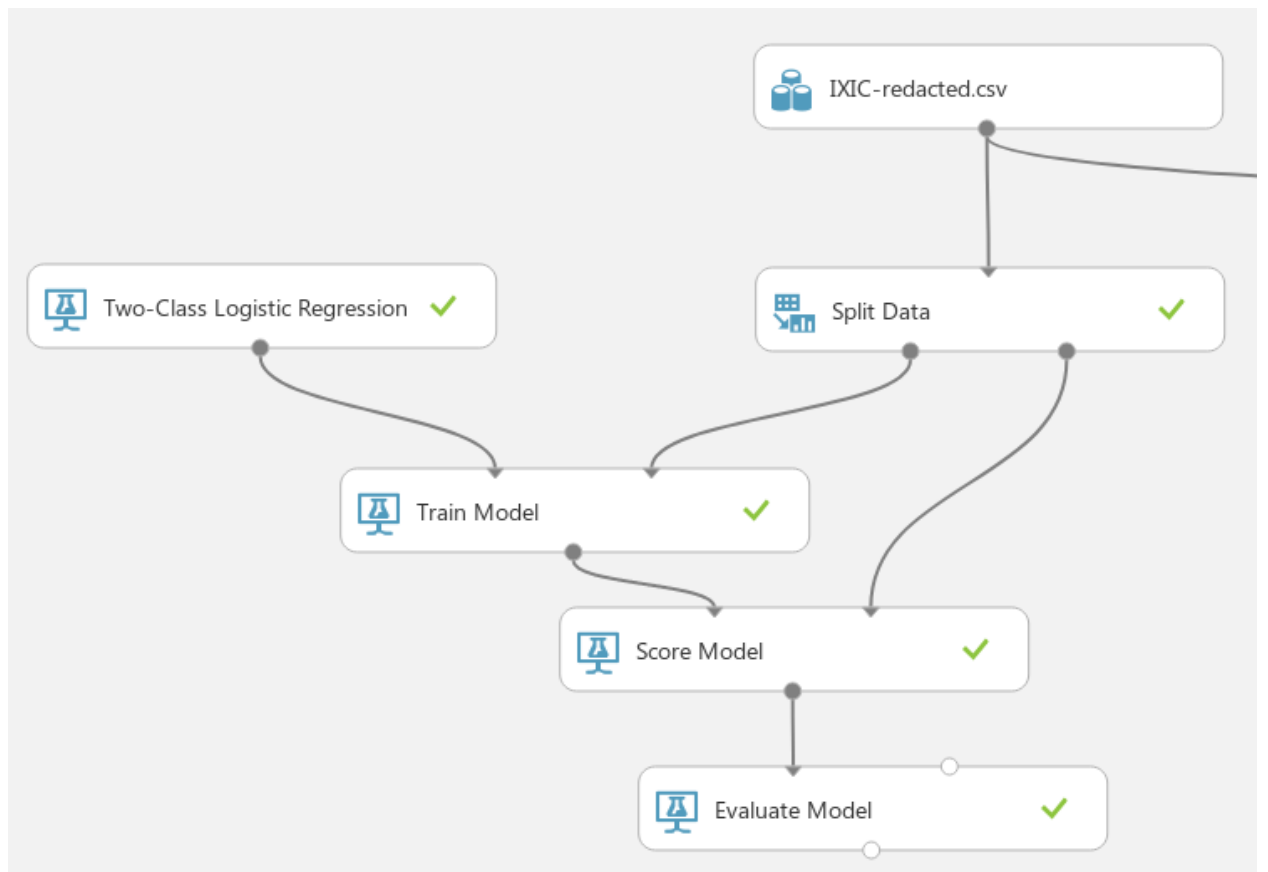
Close



В целом, результаты довольно точны, однако присутствует небольшое количество значительных отклонений.

Two-class Logistic Regression

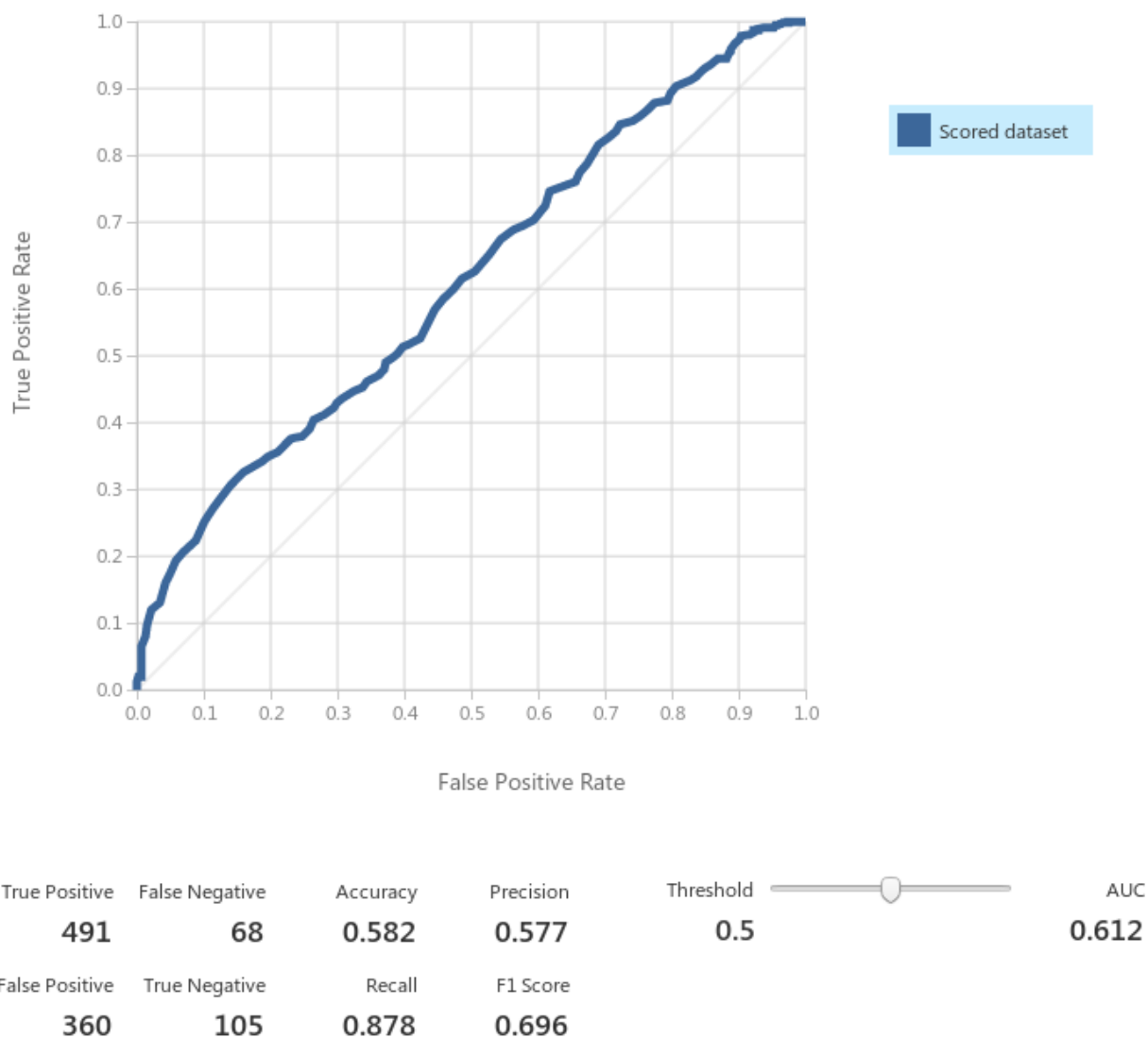
Логистическая регрессия используется для прогнозирования вероятности возникновения некоторого события путём подгонки данных к логистической кривой, тем самым подходит для классификации данного датасета.



Результаты:

Классификация прибылей акций > Evaluate Model > Evaluation results

ROC PRECISION/RECALL LIFT



Выводы.

Выполнив лабораторную работу, я ознакомилась с Microsoft Azure Machine Learning Studio. Различные алгоритмы машинного обучения очень удобно запускать в облаке. Также при работе получают наглядные модели.