

**Московский авиационный институт  
(Национальный исследовательский университет)**

Факультет: «Информационные технологии и прикладная математика»  
Кафедра: 806 «Вычислительная математика и программирование»  
Дисциплина: «Искусственный интеллект»

**Лабораторная работа № 1**  
Тема: Azure ML

Студент: Колесса Е.А.  
Группа: М80-304Б

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.

---

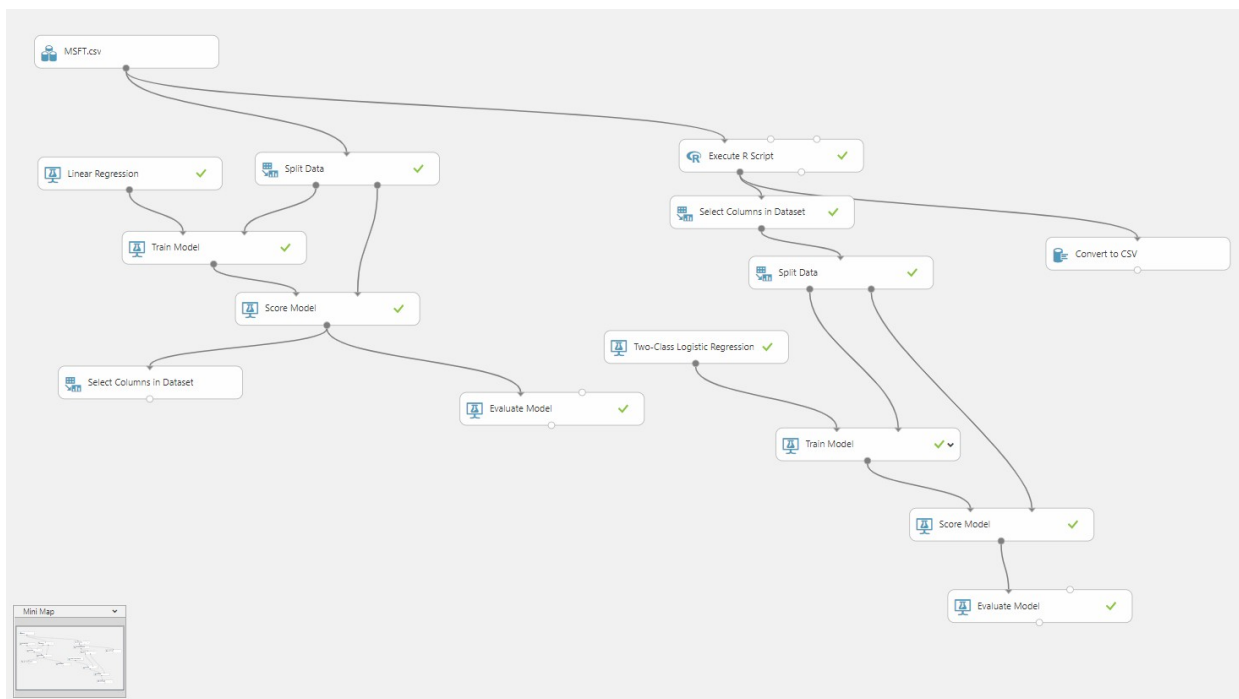
### Задача 1

---

Оценка стоимости закрытия акций с помощью регрессии. Будем использовать линейную регрессию, так как в прошлой ЛР удалось обнаружить линейную зависимость между атрибутами акций.

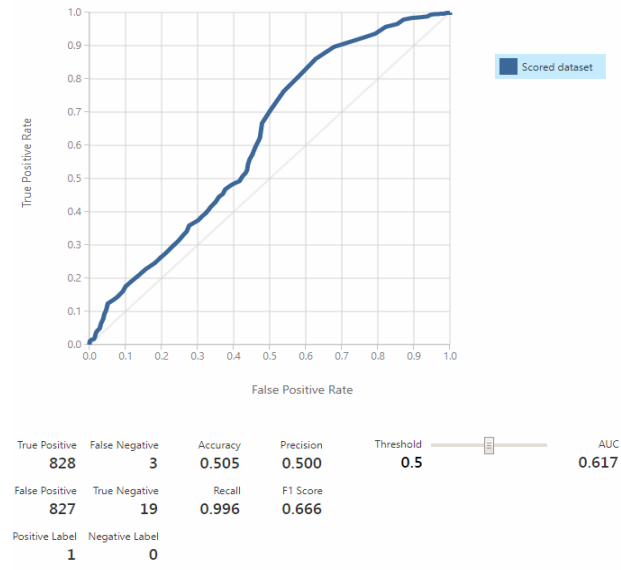
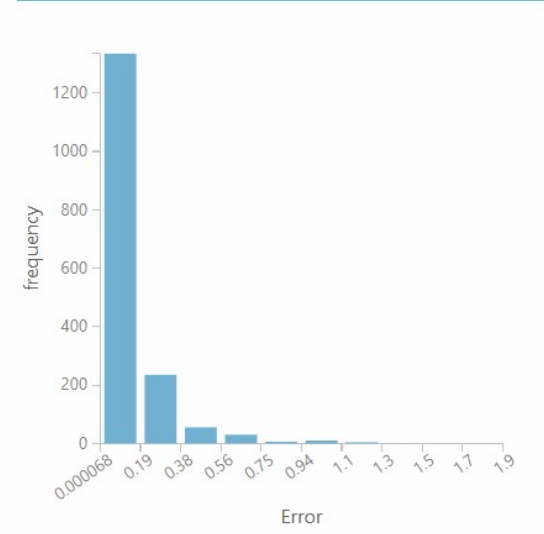
На основе имеющихся данных (дата, цена открытия, цена закрытия, max цена, min цена, скорректированная цена, объем) обучим модель определять цену закрытия. Соотношения данных на обучение/проверку = 80/20.

### Модель эксперимента



# Визуализация

## Error Histogram



Score Bin	Positive Examples	Negative Examples	Fraction Above Threshold	Accuracy	F1 Score	Precision	Recall	Negative Precision	Negative Recall	Cumulative AUC
(0.900,1.000]	0	0	0.000	0.504	0.000	1.000	0.000	0.504	1.000	0.000
(0.800,0.900]	0	0	0.000	0.504	0.000	1.000	0.000	0.504	1.000	0.000
(0.700,0.800]	0	0	0.000	0.504	0.000	1.000	0.000	0.504	1.000	0.000
(0.600,0.700]	0	0	0.000	0.504	0.000	1.000	0.000	0.504	1.000	0.000
(0.500,0.600]	828	827	0.987	0.505	0.666	0.500	0.996	0.864	0.022	0.595
(0.400,0.500]	3	19	1.000	0.496	0.663	0.496	1.000	1.000	0.000	0.617
(0.300,0.400]	0	0	1.000	0.496	0.663	0.496	1.000	1.000	0.000	0.617
(0.200,0.300]	0	0	1.000	0.496	0.663	0.496	1.000	1.000	0.000	0.617
(0.100,0.200]	0	0	1.000	0.496	0.663	0.496	1.000	1.000	0.000	0.617
(0.000,0.100]	0	0	1.000	0.496	0.663	0.496	1.000	1.000	0.000	0.617

## Metrics

Mean Absolute Error	0.121282
Root Mean Squared Error	0.211833
Relative Absolute Error	0.007049
Relative Squared Error	0.000073
Coefficient of Determination	0.999927

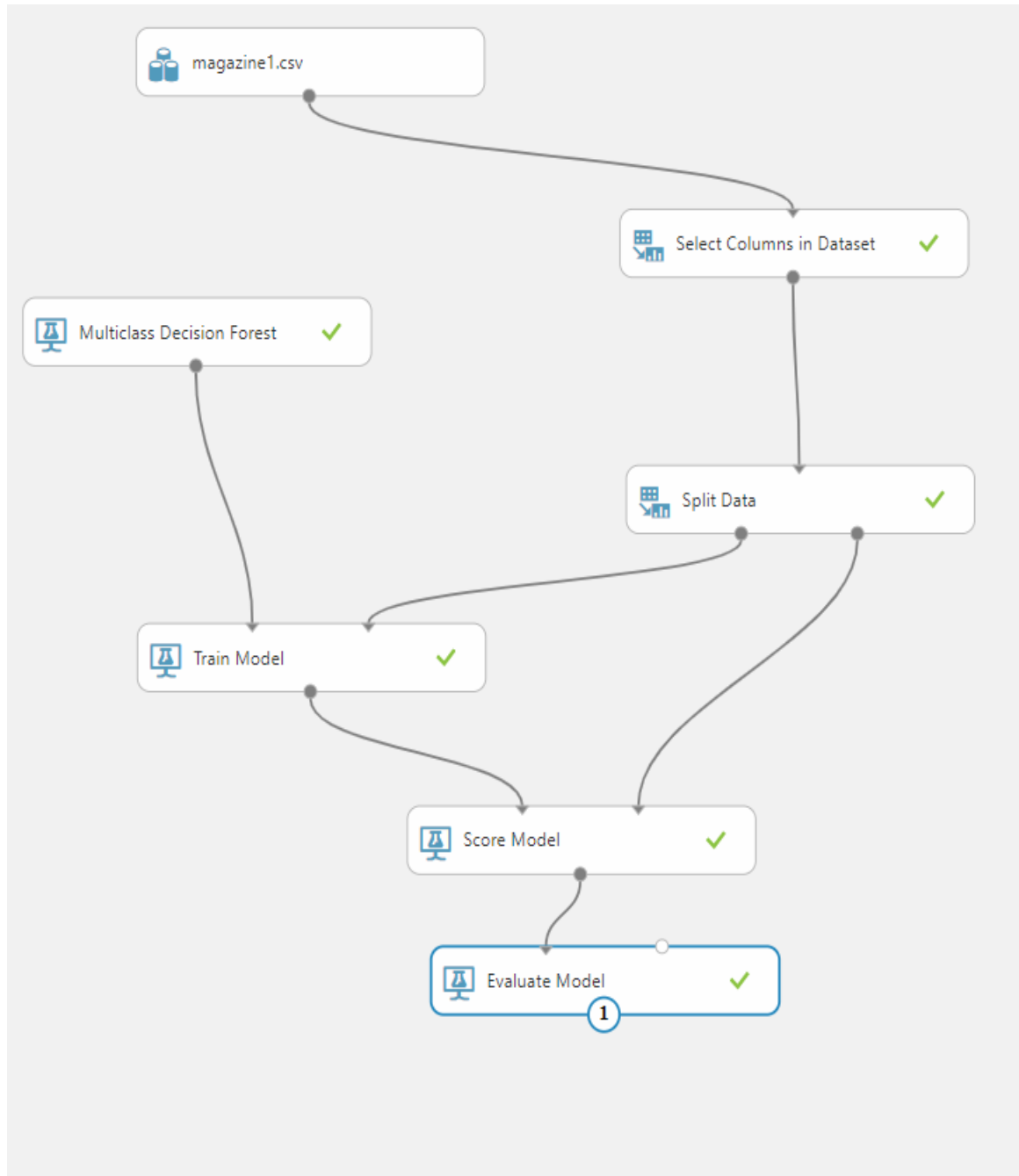
---

## Задача 2

---

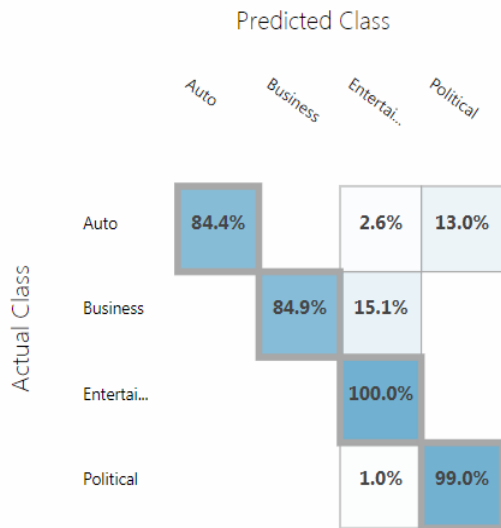
Классификация новостных статей по тематике. Для этого воспользуемся мультиклассовым алгоритмом(так как 4 тематики)

### Модель эксперимента



## Матрица ошибок и метрики

### Confusion Matrix



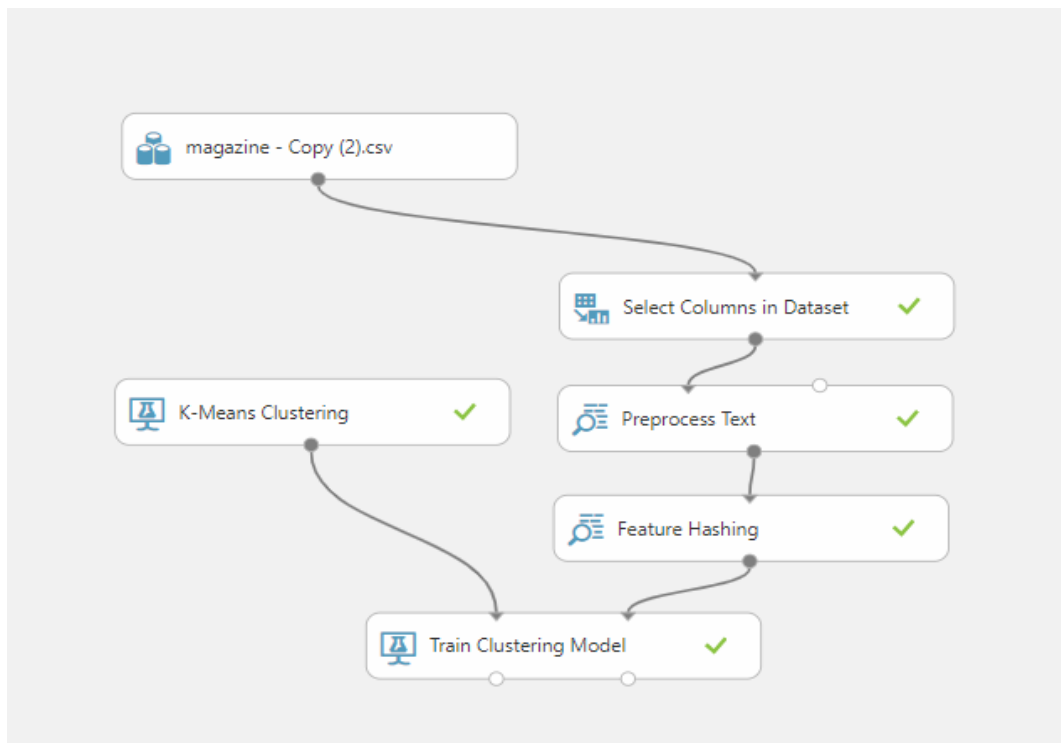
### Metrics

Overall accuracy	0.925
Average accuracy	0.9625
Micro-averaged precision	0.925
Macro-averaged precision	0.937937
Micro-averaged recall	0.925
Macro-averaged recall	0.920765

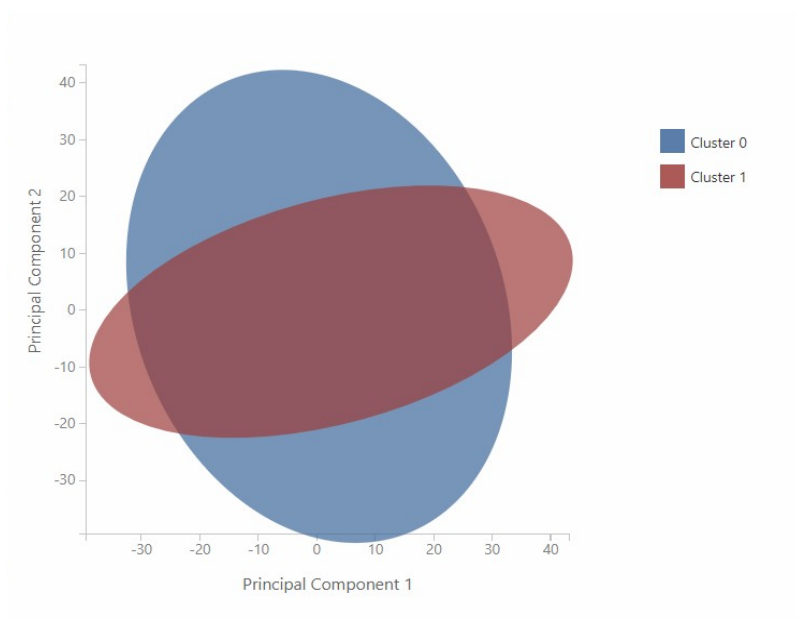
### Задача 3

Кластеризация статей и тематики.

### Модель эксперимента



## Результат обучения



---

### Выводы

---

Приятное новое знакомство. Мощно. Изячно. Мне понравилось.