**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 machine learning

Студент: Шевчук П.В.

Группа: М80-304Б

Преподаватель: Ахмед Самир Халид

Москва, 2019

**1. Постановка задачи**

Ваша задача познакомиться с платформой Azure Machine Learning, реализовав полный цикл разработки решения задачи машинного обучения, использовав три различных алгоритма, реализованные на этой платформе.

**2. Требования**

1) Уникальность решения

2) Обоснованность выбора той или иной операции

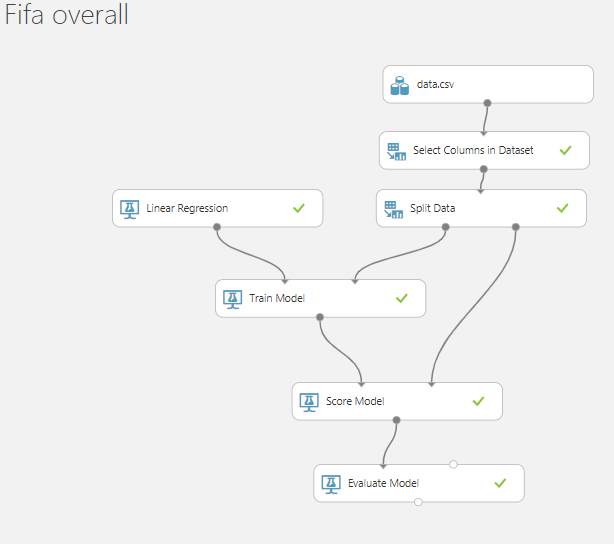
3) В отчете должны быть указаны алгоритмы, которые применялись, результаты применения этих алгоритмов, а также скрины некоторых этапов обработки данных

**3. Решение задачи**

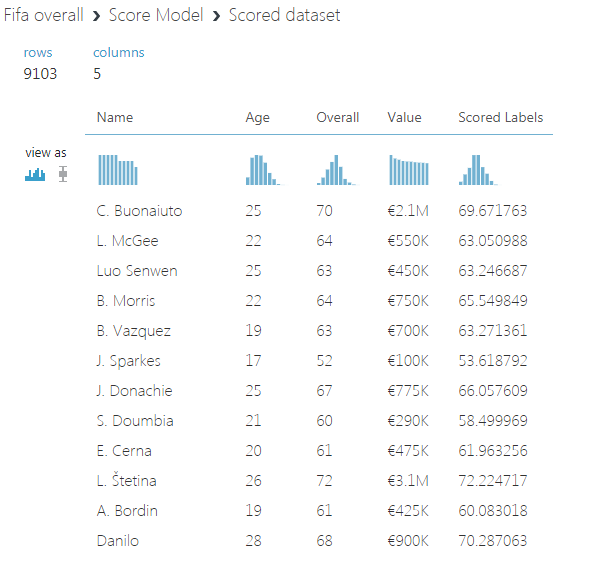
**Датасет:** <https://www.kaggle.com/karangadiya/fifa19>, данные об игроках FIFA 19

**1 алгоритм:**  Определить рейтинг игрока с помощью его данных о возрасте и цене.

Выбираем интересующие поля: рейтинг, цена, возраст и имя игрока (для удобства). Для обучения алгоритма воспользуемся классом линейной регрессии, т.к. рейтинг почти линейно зависит от цены. Подадим 80% данных на обучение, а остальные 20% на проверку. В результате видим, что предсказанные значения не сильно отличаются от действительного результата.

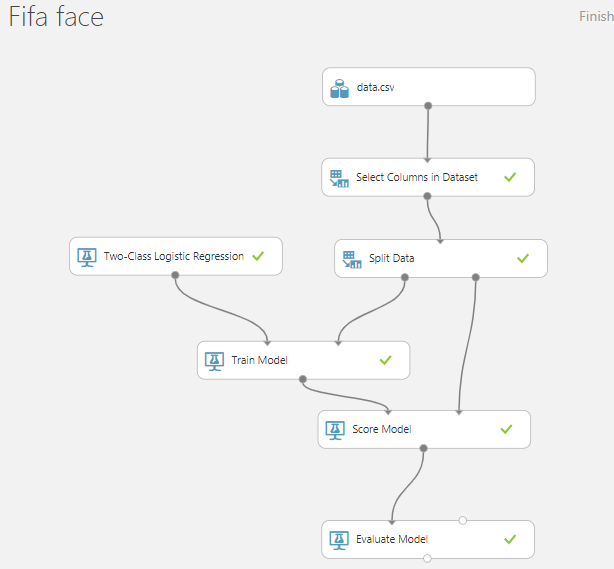


.



**2 алгоритм:**  Определить, есть ли gameface у игрока с помощью его данных о возрасте, цене.

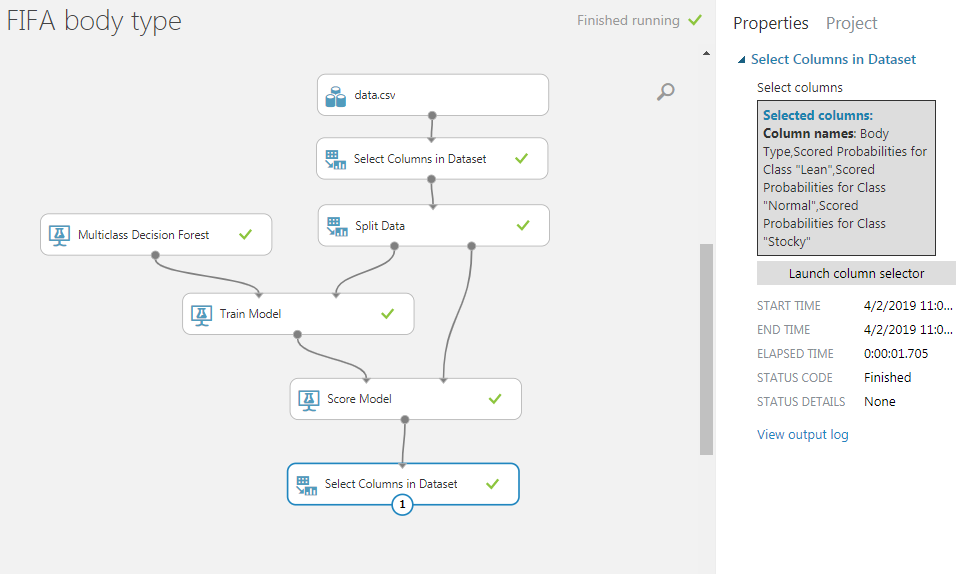
Надо предсказать категориальный признак, состоящий из двух вариантов: “Да” и “Нет”. Для этого воспользуемся методом two-class logistic regression, т.к. опять предполагается линейная зависимость от рейтинга. В результате мы видим, что большинство игроков не имеют gameface, но есть игроки с не самым большим рейтингом, которые имеют его. Это может быть связано с историей игрока или с другими причинами, поэтому алгоритм не может в точности дать верный результат при исходных данных.

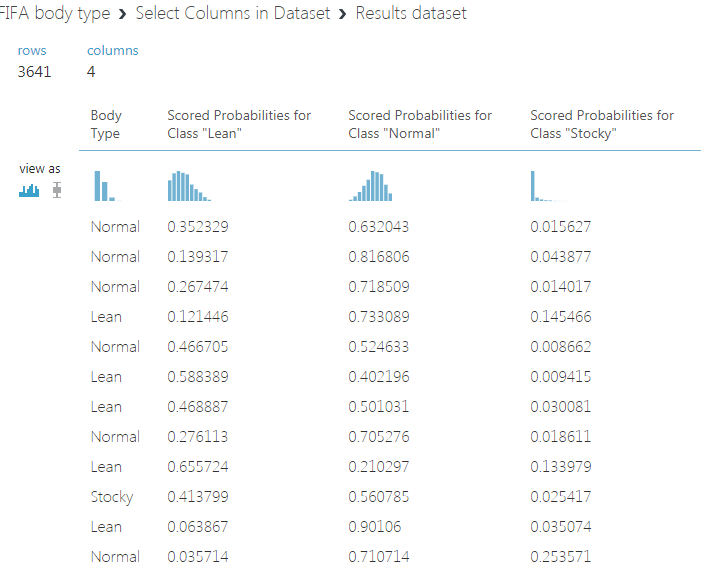


### 

### 3 алгоритм: Определить, какая вероятность того или иного телосложения у игрока.

Выбираем поля: возраст, рейтинг, рост, вес и вид телосложения. Так как категориальный признак имеет 3 варианта, то воспользуемся методом multiclass decision forest. Этот метод быстро изучает машину и работает с высокой точностью. Выведем признаки, относящиеся только к типу телосложения. Наблюдаем вероятность одного из трёх видов, которые может получить игрок.





### 4. Вывод

В ходе данной лабораторной работы освоен алгоритм работы с Azure machine learning. С помощью этого облачного сервиса я научился прогнозировать разные признаки. Надеюсь, эта система поможет мне для дальнейшей работы в машинном обучении.