**1. Загрузил train и val datasets**

**2. Воспользовался функциями, для получения информации о данных хранящихся в**

**датасетах**

**df.head()**

**df.describe()  
df.dtypes**

**3. Поскольку присутствуют na значения воспользуемся функцией Fill\_na() которая замещает значения na на значения среднее по столбцу**

**4. Перетасовал строки в dataframe с помощью функции shuffle, поскольку все строки упорядочены по ‘class’**

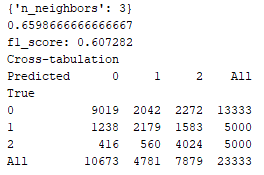
**5. Далее я удалил столбцы в train, val dataframe, которые не должны присутствовать, либо негативно влияют на результат классификации**

**5. Нормализовал данные, для приведения значений во всех столбцах к единой шкале.**

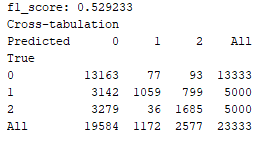
**6. Применил различные методы классификации. Результатам показали, что нейронная сеть имеет наибольшую точность классификации объектов.**

**Точность классификации различных методов:**

1. **KNN GridSearchCV**



1. **Random forest criterion = ‘entropy’**

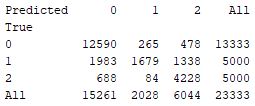


1. **Neural network**

**Training set score: 0.781500**

**Test set score: 0.792740**

**F1\_score 0.7080240034253177**



**7.Результаты классификации занес в файл results\turchin**