**Московский государственный технический университет им. Н.Э.Баумана**

**Кафедра «Системы обработки информации и управления»**

Рубежный контроль №2 по дисциплине

«Методы машинного обучения»

Выполнил:

Студент группы ИУ5И-21М Ся Бэйбэй

Москва — 2022г.

# Решение задачи классификации текстов.

Необходимо решить задачу классификации текстов на основе любого выбранного Вами датасета (кроме примера, который рассматривался в лекции). Классификация может быть бинарной или многоклассовой. Целевой признак из выбранного Вами датасета может иметь любой физический смысл, примером является задача анализа тональности текста.

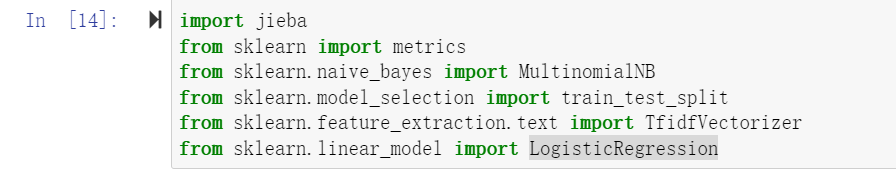
Необходимо сформировать два варианта векторизации признаков - на основе CountVectorizer и на основе TfidfVectorizer.

В качестве классификаторов необходимо использовать два классификатора по варианту для Вашей группы:

| **Группа** | **Классификатор №1** | **Классификатор №2** |
| --- | --- | --- |
| ИУ5-21М, ИУ5И-21М | [LogisticRegression](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html) | [Multinomial Naive Bayes - MNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html) |
| ИУ5-22М, ИУ5И-22М | [RandomForestClassifier](https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html) | [Complement Naive Bayes - CNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.ComplementNB.html) |
| ИУ5-23М, ИУ5И-23М | [LinearSVC](https://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html) | [Multinomial Naive Bayes - MNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html) |
| ИУ5-24М, ИУ5И-24М | [KNeighborsClassifier](https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html) | [Complement Naive Bayes - CNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.ComplementNB.html) |

Мой вариант: [LogisticRegression](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html)&[Multinomial Naive Bayes - MNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html)

**Список Библитека:**

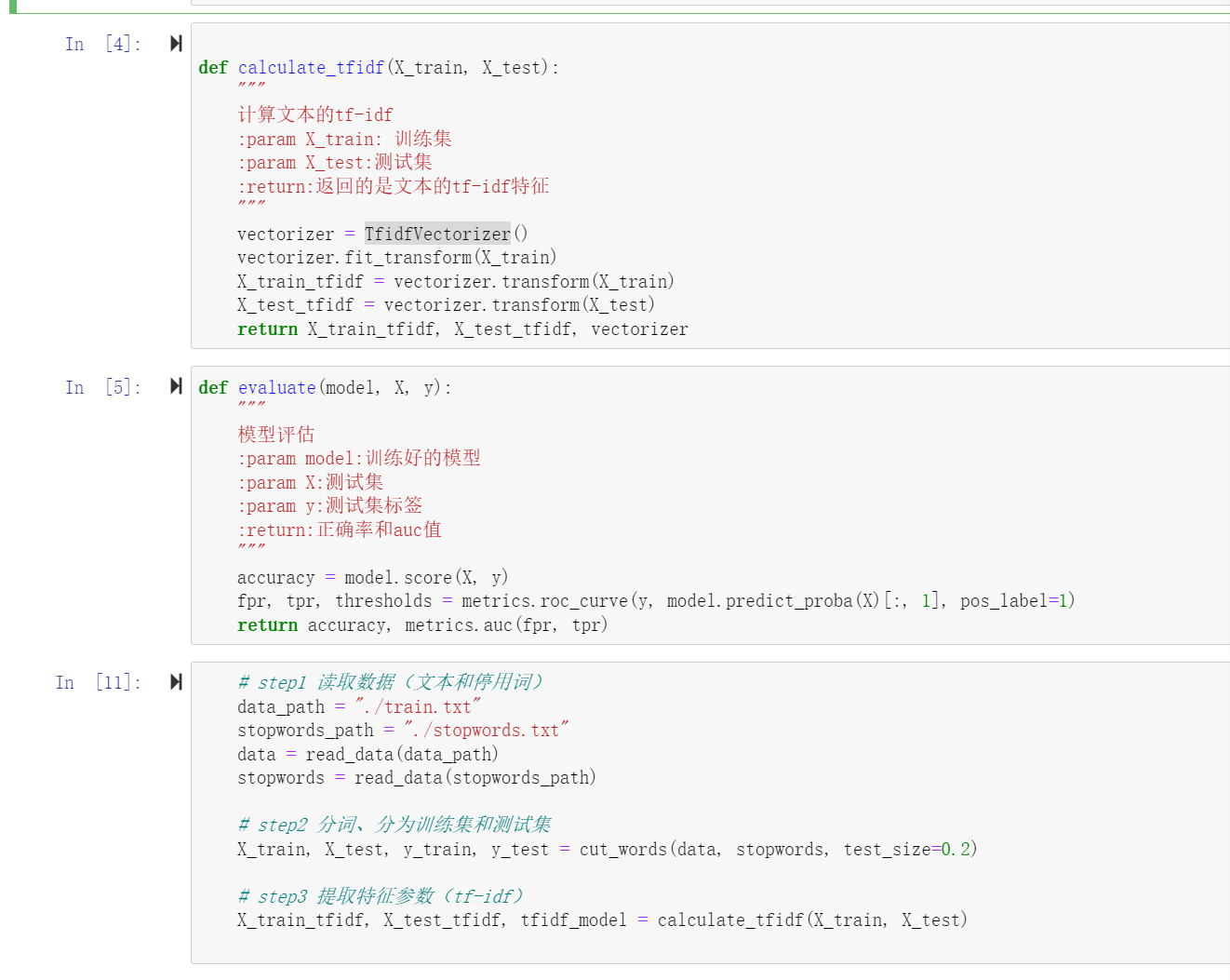
****

**читать данные:**

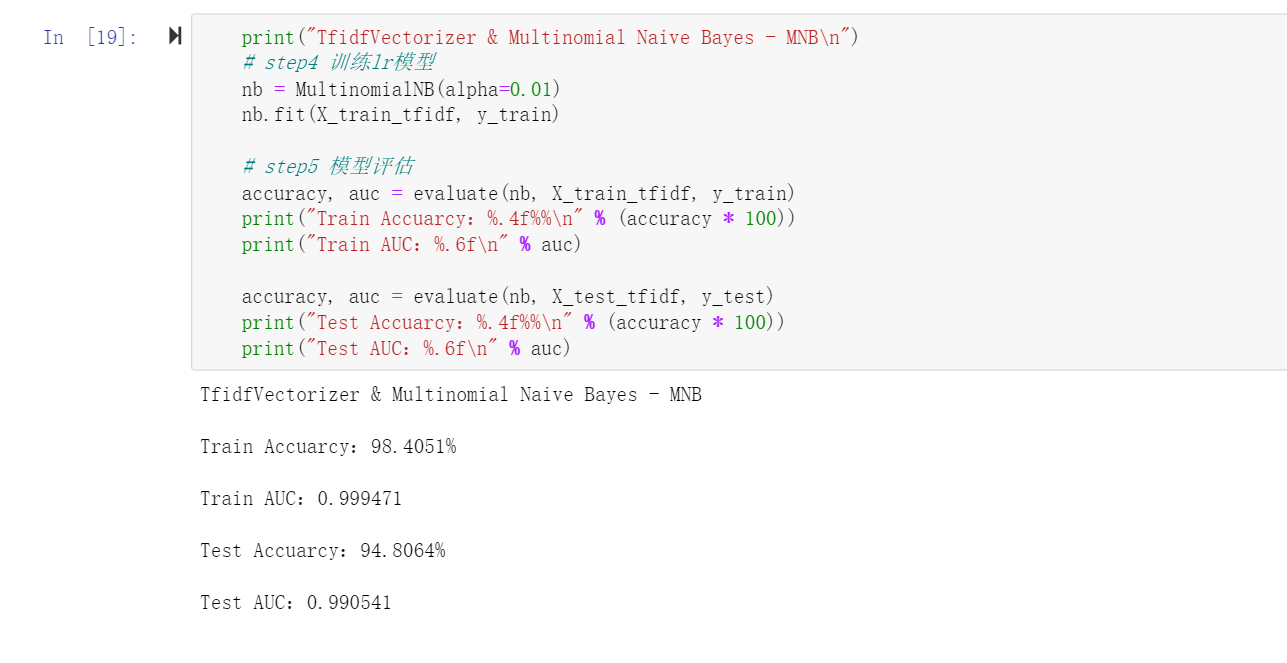
****

**Разделите набор данных на обучающий и тестовый наборыЖ**

****

****

[LogisticRegression](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html)&[Multinomial Naive Bayes - MNB](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html) :

****

****