



Search or jump to...



[Pull requests](#) [Issues](#) [Marketplace](#) [Explore](#)



ugapanyuk / **ml_course_2022**

Public

Watch 3

Fork 5

Star 7

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#)

TMO_RK_2

ugapanyuk edited this page 11 days ago · 3 revisions

Рубежный контроль №2

Необходимо подготовить отчет по рубежному контролю и разместить его в Вашем репозитории. Вы можете использовать титульный лист, или в начале ноутбука в текстовой ячейке указать Ваши Ф.И.О. и группу.

Тема: Методы построения моделей машинного обучения.

Задание. Для заданного набора данных (по Вашему варианту) постройте модели классификации или регрессии (в зависимости от конкретной задачи, рассматриваемой в наборе данных). Для построения моделей используйте методы 1 и 2 (по варианту для Вашей группы). Оцените качество моделей на основе подходящих метрик качества (не менее двух метрик). Какие метрики качества Вы использовали и почему? Какие выводы Вы можете сделать о качестве построенных моделей? Для построения моделей необходимо выполнить требуемую предобработку данных: заполнение пропусков, кодирование категориальных признаков, и т.д.

- Для студентов групп ИУ5-61Б, ИУ5-62Б, ИУ5-63Б, ИУ5-64Б, ИУ5-65Б, РТ5-61Б номер варианта = номер в списке группы.
- Для студентов групп ИУ5Ц-81Б, ИУ5Ц-82Б, ИУ5Ц-83Б, ИУ5Ц-84Б номер варианта = 25 + номер в списке группы.
- Для студентов группы ИУ5И-61Б номер варианта = 30 + номер в списке группы.
- При решении задач можно выбирать любое подмножество признаков из приведенного набора данных.
- Для сокращения времени построения моделей можно использовать фрагмент набора данных (например, первые 200-500 строк).
- Методы 1 и 2 для каждой группы приведены в следующей таблице:

Группа	Метод №1	Метод №2
ИУ5-61Б, ИУ5Ц-81Б, ИУ5И-61Б	Линейная/логистическая регрессия	Случайный лес
ИУ5-62Б, ИУ5Ц-82Б	Метод опорных векторов	Случайный лес
ИУ5-63Б, ИУ5Ц-83Б	Дерево решений	Случайный лес
ИУ5-64Б, ИУ5Ц-84Б	Линейная/логистическая регрессия	Градиентный бустинг
ИУ5-65Б	Метод опорных векторов	Градиентный бустинг
РТ5-61Б	Дерево решений	Градиентный бустинг

- Варианты наборов данных приведены в следующем списке:

1. https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_boston.html#sklearn.datasets.load_boston
2. https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_iris.html#sklearn.datasets.load_iris
3. https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_wine.html#sklearn.datasets.load_wine
4. <https://www.kaggle.com/carlolepelaars/toy-dataset>
5. <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset>
6. <https://www.kaggle.com/mohansacharya/graduate-admissions> (файл Admission_Predict.csv)
7. <https://www.kaggle.com/mohansacharya/graduate-admissions> (файл Admission_Predict_Ver1.1.csv)
8. <https://www.kaggle.com/lava18/google-play-store-apps>
9. <https://www.kaggle.com/rubenssjr/brasilian-houses-to-rent>
10. <https://www.kaggle.com/jessemostipak/hotel-booking-demand>
11. <https://www.kaggle.com/karangadiya/fifa19>
12. <https://www.kaggle.com/fivethirtyeight/fivethirtyeight-comic-characters-dataset> (файл dc-wikia-data.csv)
13. <https://www.kaggle.com/fivethirtyeight/fivethirtyeight-comic-characters-dataset> (файл marvel-wikia-data.csv)
14. <https://www.kaggle.com/noriuk/us-education-datasets-unification-project> (файл states_all.csv)
15. <https://www.kaggle.com/noriuk/us-education-datasets-unification-project> (файл states_all_extended.csv)
16. <https://www.kaggle.com/san-francisco/sf-restaurant-scores-lives-standard>
17. <https://www.kaggle.com/mathan/fifa-2018-match-statistics>
18. <https://www.kaggle.com/rhuebner/human-resources-data-set>
19. <https://www.kaggle.com/arindam235/startup-investments-crunchbase>
20. <https://www.kaggle.com/fmejia21/trump-impeachment-polls>
21. <https://www.kaggle.com/oreojam/formula-e-world-championship-race-results>
22. <https://www.kaggle.com/atmcfarland/historical-us-president-physical-data-more>
23. <https://www.kaggle.com/ajaypalsinghlo/world-happiness-report-2021>
24. <https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease>
25. <https://www.kaggle.com/prachi13/customer-analytics>
26. <https://www.kaggle.com/brsdincer/star-type-classification>
27. <https://www.kaggle.com/fedesoriano/company-bankruptcy-prediction>
28. <https://www.kaggle.com/roysouravcu/forbes-billionaires-of-2021>
29. <https://www.kaggle.com/dhruvildave/currency-exchange-rates>
30. <https://www.kaggle.com/paytonfisher/sp-500-companies-with-financial-information>
31. <https://www.kaggle.com/kmldas/loan-default-prediction>
32. <https://www.kaggle.com/rashikrahmanpritom/disney-movies-19372016-total-gross>

▼ Pages 23

Find a Page...

▶ Home

▶ COURSE_MAD

▶ COURSE_MMO

▶ COURSE_TMO

▶ DSLIST

▶ DZ_MMO

▶ IDE

▶ LAB_MMO_DATA_STORY

▶ LAB_MMO_FEATURES

▶ LAB_MMO_FEATURES_2

▶ LAB_MMO_RECOM

▶ LAB_TMO_EDA_VISUALIZATION

▶ LAB_TMO_KNN

▶ LAB_TMO_MISSING

▶ LAB_TMO_ENSEMBLES

Show 8 more pages...

Clone this wiki locally

https://github.com/ugapanyuk/ml_



© 2022 GitHub, Inc.

[Terms](#)

[Privacy](#)

[Security](#)

[Status](#)

[Docs](#)

[Contact GitHub](#)

[Pricing](#)

[API](#)

[Training](#)

[Blog](#)

[About](#)