**Задание к экзамену.**

Построить модель ML в виде практической Jupyter-тетрадки

Написать и оформить как тьюториал по используемым методам в решаемой задаче в виде второй теоретической Jupyter-тетрадки..

То есть определения, методы, алгоритмы, метрики, используемая литература, ссылки

Максимально визуализировать решения, прописать выводы.

Выложить на colab, github.

Если используете задачу с kaggle, то выкладываем результаты соревнования

Этапы работы и содержание:

1. Собрать данные, проанализировать, исправить и т.п.
2. Произвести визуальный анализ данных, используя различные библиотеки, сделать выводы о закономерностях, лежащих внутри этих данных.
3. Преобразовать и стандартизировать входные данные.
4. Проанализировать распределения переменных модели.
5. Выбрать методы анализа данных.
6. Загрузить данные и разделить данные для обучения.
7. Реализовать алгоритмы машинного обучения.
8. Получить ROC кривые для вариантов без нормировки.
9. Получить ROC кривые для вариантов с нормировкой признаков.
10. Получить другие оценки моделей.
11. Визуализируйте методы понижения размерности.
12. Разработать интерактивную панель мониторинга или панель визуализации результатов модели, обосновать выбор инструментов.
13. др.

Можно собрать данные из разных источников, можно развернуть модель.

Примеры данных:  характеристики банков и их присутствие их клиентских приложений в AppStore и PlayMarket , данные о продажах табачной продукции за последний год, временной ряд курсов акций, валют и др.

Варианты проектов

1. Страховая компания хочет улучшить прогнозирование денежных потоков за счет более точного прогнозирования расходов на пациента с использованием демографических и основных показателей риска для здоровья пациентов во время госпитализации.  Создать веб-приложение, в котором демографические данные и информация о здоровье пациента вводятся в веб-форму для прогнозирования расходов .
2. Разработать интерактивную панель мониторинга в режиме реального времени для финансовых данных компании. Использовать в аналитических панелях разнообразные визуальные элементы: графики, столбчатые диаграммы, круговые, пузырьковые, блочные, диаграммы воронок и др.
3. *Анализ связей заемщиков с другими юридическими и физическими лицами. Написать алгоритм построения графа связей между юридическими лицами.* анализ связей осуществляется методами графовой аналитики, которая позволяет автоматизировать поиск и обновление связей по клиентам банка. Помимо этого, автоматизация поиска связей между компаниями позволяет рассчитывать числовые метрики, описывающие эти связи, которые в дальнейшем могут быть использованы и для решения других задач, в частности для скоринга корпоративных клиентов. Рассчитать метрики: «Количество соседей», «Доля компаний с низким рейтингом», «Доля компаний-банкротов», «Доля закрытых компаний», «Доля новообразованных компаний» , дать характеристику, оптимальный размер.

и мн другое

|  |  |
| --- | --- |
| Проектная группа (Фамилии) | Тема |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |