

Здесь разная информация и всякие мысли,
которые у меня накапливаются по ML,
математике и иногда по физике.

Содержание

1	Математика	2
1.1	Статистика и теорвер	2
1.2	Принципы работы разных моделей	2
1.2.1	НОРМАЛИЗАЦИЯ - for what???	2
1.3	Оценка качества моделей	2
2	Прога	2
2.1	Мысли про примитивные либы(пандас, склерн и тп)	2
2.2	Бустинг в catboost (и не только)	2
2.3	Заметки о pytorch	2
2.4	Здесь будет все про разные архитектуры сетей	2
2.5	О железе для МЛ	2
2.6	О jit компиляции в питоне и оптимизации вычислений в нем разными пакетами	2
2.7	* О параллельных вычислениях	2
3	*О базах данных	2
4	*Заметки по поводу Julia	3
5	**Физика???	3
6	***Не МЛьная математика?????	3

1 Математика

1.1 Статистика и теорвер

<http://npm.mipt.ru/books/lab-intro/Ch3.html>

- про метод макс правдоподобия, хи квадрат и МНК

https://drive.google.com/file/d/17h2YiMHxdlnVp_x9_QAfeVhssiiASQE/view?usp=sharing

- statistics for machine learning

https://drive.google.com/file/d/17h2YiMHxdlnVp_x9_QAfeVhssiiASQE/view?usp=sharing

- ссылка на мой summary теорвера для МЛ

Мои замечания всякие:

1.2 Принципы работы разных моделей

1.2.1 НОРМАЛИЗАЦИЯ - for what???

1.3 Оценка качества моделей

https://drive.google.com/file/d/1yIMdm0YVFaqLo6qkMitDXFquPbyo9_sB/view?usp=sharing

- статья с архива про оценку качества, все очень подробно

2 Прога

2.1 Мысли про примитивные либы(пандас, склерн и тп)

2.2 Бустинг в catboost (и не только)

<https://habr.com/ru/company/otus/blog/527554/>

- супер мега статья на хабре про кэтбуст.

https://github.com/Alphonsce/ML_learning/blob/main/house_prices/tree_model.ipynb

- Также см мой код на гитхабе (здесь топ 10% в house prices на бустинге)

2.3 Заметки о pytorch

2.4 Здесь будет все про разные архитектуры сетей

2.5 О железе для МЛ

2.6 О jit компиляции в питоне и оптимизации вычислений в нем разными пакетами

Есть Cython - писать в питоне сишный код

Есть Numba - она при помощи менеджера @jit (контекстный менеджер - это по-сути функция принимающая функцию) позволяет некоторые вещи jit компилировать внутри функции.

Лучше использовать @njit, тогда будет использоваться только то, что оптимизировано и сваливаться с ошибкой если перешло к обычному питону.

2.7 * О параллельных вычислениях

3 *О базах данных

Реляционная БД - это набор отношений - то есть табличка по-сути, в ней нельзя сделать вложенность.

Штуки типо Postgresql и пр - это СУБД, у каждой СУБД +- реализован набор базисных операций, которые умеет делать язык запросов SQL

- 4 *Заметки по поводу Julia
- 5 **Физика???
- 6 ***Не МЛьная математика?????