

Куцев Роман

4 место

1	10011000 (pavelostaaa, MrAxcel)	84489992
2	Антихайп (cszdr, vasyarv)	82909074
3	Classics(ivanicki-i, Amir14111) 	80423665
4	kutcu	79701913
5	sal.vios	78750097
6	alexisonfireyar 	75905679
7	Ясонов Евгений	75734019
8	Антон Патрикеев	75606736
9	Aiaiai (karfly, illusionww)	74950292
10	Iviconun(FireSonics, mrk.andreev)	74761301

Главные проблемы:

- Время (31 час)
- Данные (30 гб)

Времени не было, поэтому:

- Не использовал кросс-валидацию
- Только проверенные методы, которые зарекомендовали себя на других соревнованиях
- Не тюнил модели

Главная ошибка: не нашел команду

Зато призом не надо делиться ухаххаха

Данные:

- Их было много
- Очень много
- 90% данных не использовалось в финальном решении

Железо:

- Стационарный компьютер
- i5 6600 4 ядра
- Очень большой SSD
- 16 гб ОЗУ + 40 гб подкачка SSD

Больше данных = меньше участников

- 100+ участников пришли на очный этап
- 40 участников отослали хотя бы 1 посылку
- 22 участника получили решение выше бейзлайна

За день до хакатона:

Обязательно возьмите с собой рабочий ноутбук и зарядное устройство. Также, пожалуйста, убедитесь, что на вашем ноутбуке есть достаточно места: данные в сжатом виде занимают 1 Гб, а в разжатом до 30 Гб.

На хакатоне:



Nina Morenko

кто-нибудь может скинуть хороший tutorial/описание, как можно забронировать инстансы, залить свои данные на Amazon AWS и простенький пример Notebook, где уже не локально со своей оперативной памятью работаешь, а на сервере кодишь?

Решение:

- 1) Кол во пешеходов —
- 2) Кол во машин —
- 3) 25%, 60%, 75% машин +
- 4) Путь пешеходов —
- 5) Расчеты стоимости ходов —
- 6) Кол от лодок —
- 7) Аг регизация соседних клеток +
- 8) Мат лог эвристики +
- 9) Измерения приходов ряды —
- 8) Итерации с соседними клетками +
- 9) Анализ алгоритмов — НЕ УСПЕЮ

P_0	P_1	P_2
P_3	P	P_5
P_6	P_4	P_7



私は日本語を勉強します。

- Решение состояло из бейзлайна организаторов и кусков кода из других задач
- Не использовал данные о сотовой связи
- В основном использовал данные метеостанций Netatmo. Mean, std, min, max, 25%, 75%.
- Удалил данные о сотовой связи, затем удалил дубликаты, dataset сократился в 3.5 раза
- Добавил фичи: кол-во пешеходов, кол-во машин

Получил 0.78

«Стекинг»:

p	p0	p1	p2	p3	p5	p6	p7	p8	mean	min	sum	max
0.114295	0.069358	0.086025	0.021341	0.087581	0.093450	0.166121	0.020389	0.110514	0.085453	0.020389	0.874917	0.874917
0.738168	0.098092	0.765430	0.097355	0.097397	0.097447	0.396871	0.096763	0.411669	0.578034	0.396871	3.287043	3.287043
0.111346	0.098092	0.096440	0.097355	0.192703	0.097447	0.097518	0.078358	0.105630	0.122009	0.078358	0.688403	0.688403
0.084067	0.098092	0.103149	0.108659	0.084037	0.089303	0.060465	0.109839	0.098508	0.091360	0.060465	0.791344	0.791344
0.029331	0.098092	0.096440	0.097355	0.097397	0.097447	0.097518	0.096763	0.098508	0.029331	0.029331	0.087994	0.087994

С помощью Kfold предсказываем train, предсказываем test. Находим и добавляем предсказания соседних клеток. Финальный сабмит: нейронная сеть на новых фичах.

Получил 0.79

