

## Оглавление

Медицина .....	2
Григорий Сапунов gs@inten.to – курс по МЛ в медицине – .....	2
Опыт создания регистра по флебологии – .....	2
Диссертация на тему «Вероятностная оптимизация выводов и реализация на её основе методов, .....	2
Метод оценки значений параметров математической модели медицинской диагностики – .....	2
Другое .....	2
BEST KAGGLE KERNELS .....	3
SQLITE links .....	5
ML links .....	5
Видео участия в соревнованиях .....	6
Kaggle Santander Customer Transaction Prediction — Василий Рязанов .....	6
Дайджест соревнований по анализу данных – Сергей Брянский - transactions .....	6
Kaggle Allstate Claims Severity: предсказание тяжести страхового случая — Алексей Носков .....	6
Kaggle Home Credit: определение риска дефолта по кредитам — Евгений Патеха .....	7
Tinkoff Data Science Challenge: анализ задачи предсказания выбора кредита — Алексей Чернобровов ..	7
ML Boot Camp III: предсказание ухода пользователей из онлайн-игры — Михаил Карачун .....	7
Kaggle Two Sigma Financial Modeling Challenge: предсказание финансовых величин — С. Семёнов .....	7
Kaggle Corporación Favorita Grocery: предсказание продаж сети супермаркетов — Андрей Филимонов .	7
Определение дома и работы клиента по банковским транзакциям – Сергей Гайдаенко .....	7
OpenAI Retro: как казаки retro контекст решали — Сергей Колесников .....	7
Kaggle Santander: предсказание размера транзакций потенциальных клиентов — Анатолий Ильенков .	7
Kaggle Mercedes Benz: предсказание времени тестирования автомобилей — Данила Савенков .....	7
Чернобровов Подход к автоматизации решений задач ML на примере HackerEarth ML .....	7

## Медицина

### Григорий Сапунов [gs@inten.to](mailto:gs@inten.to) – курс по МЛ в медицине –

Пока можно подключиться к слаку

курса: [https://join.slack.com/t/deeplearning4biology/shared\\_invite/enQtMjU2MDM2NDA2Nzg5LTFhMTcyNTVIN2E5ODA4YmNIOTZiNTQ1YWQyYTIiNmI5MDVjNDMzOWMxMjg5ZjYxNDg4ZGNjNDJmYjFmZTEwYjY](https://join.slack.com/t/deeplearning4biology/shared_invite/enQtMjU2MDM2NDA2Nzg5LTFhMTcyNTVIN2E5ODA4YmNIOTZiNTQ1YWQyYTIiNmI5MDVjNDMzOWMxMjg5ZjYxNDg4ZGNjNDJmYjFmZTEwYjY)

Материалы курса (включая видео) выкладываются в гитхаб: <https://github.com/che-shr-cat/deep-learning-for-biology-hse-2019-course>

### Опыт создания регистра по флебологии –

<https://mibsnews.ru/kak-zapustit-registr-patsientov-s-nulya/>

### Диссертация на тему «Вероятностная оптимизация выводов и реализация на её основе методов,

алгоритмов и программных комплексов автоматизированной диагностики наследственных болезней», скачать бесплатно автореферат по специальности ВАК РФ 05.13.18 - Математическое моделирование, численные методы и комплексы программ -

<https://www.dissercat.com/content/veroyatnostnaya-optimizatsiya-vyvodov-i-realizatsiya-na-ee-osnove-metodov-algoritmov-i-progr>

### Метод оценки значений параметров математической модели медицинской диагностики –

<http://tekhnosfera.com/view/371806/d?#?page=2>

## Другое

Создание и чтение QR-кода на Python в 3 минуты (OpenCV и qrcode) / Хабр

– <https://m.habr.com/ru/post/468865/>

Beautiful Machine Learning Pipeline with Scikit-Learn | Codementor –

<https://www.codementor.io/bruce3557/beautiful-machine-learning-pipeline-with-scikit-learn-uiqapbxuj>

Курс по машинному обучению на Coursera от Яндекса и ВШЭ / Яндекс corporate blog / Хабр -

<https://habr.com/en/company/yandex/blog/269175/>

Спортивный анализ данных, или как стать специалистом по data science / Яндекс corporate blog / Хабр -

<https://habr.com/en/company/yandex/blog/327444/>

Kaggle: История о том как мы учились предсказывать релевантность поисковых запросов и заняли 3-е место

/ Хабр - [https://habr.com/en/post/305026/#comment\\_9697464](https://habr.com/en/post/305026/#comment_9697464)

## BEST KAGGLE KERNELS

<https://www.kaggle.com/kernels/scriptcontent/4523019/download>

sklearn Pipeline playground for 10+ classifiers - <https://www.kaggle.com/chmaxx/sklearn-pipeline-playground-for-10-classifiers>

A Detailed Regression Guide with House-pricing - <https://www.kaggle.com/masumrumi/a-detailed-regression-guide-with-house-pricing>

Simple House Price Prediction (good plots) - <https://www.kaggle.com/kernels/scriptcontent/21031274/download>

how write the article - Comprehensive data exploration with Python - <https://www.kaggle.com/kernels/scriptcontent/19403046/download>

работа с моделями - надо разобраться - <https://www.kaggle.com/sonnihs/house-prices>

работа с мета моделями - моя основа для этого соревнования - хорошо расписаны класс для моделей и стекинга - <https://www.kaggle.com/serigne/stacked-regressions-top-4-on-leaderboard>

статья от организаторов как готовить файл с результатами - <https://www.kaggle.com/dansbecker/submitting-from-a-kernel>

легкий анализ через модели без части фича-инжиниринга + нейронка - <https://www.kaggle.com/apapiu/regularized-linear-models>

статья от организаторов - как заполнять неизвестное и работа с вылетами - <https://www.kaggle.com/dansbecker/handling-missing-values>

статья от организаторов об xgbust и еще есть пайплайны - <https://www.kaggle.com/dansbecker/xgboost>

от организаторов про jne-hot-encoder и расчет сред абс ошибки - <https://www.kaggle.com/dansbecker/using-categorical-data-with-one-hot-encoding>

инжиниринг и развесовка по сезонам года - <https://www.kaggle.com/chmaxx/slim-data-cleaning-modelling-weighted-ensemble>

очень важно!!! - работа с данными и модели - <https://www.kaggle.com/jesucristo/1-house-prices-solution-top-1>

надо разобраться - работа с данными, хороший скор, - <https://www.kaggle.com/shaygu/house-prices-begginer-top-7>

список корреляций - <https://www.kaggle.com/poonaml/house-prices-data-exploration-and-visualisation>

ручной анализ данных - посмотреть интересные способы отлавливания ошибок -

<https://www.kaggle.com/laurenstc/top-2-of-leaderboard-advanced-fe>

настройка гиперпараметров - <https://www.kaggle.com/vikumsw/guide-on-hyperparameter-tuning-using-gridsearchcv>

мне понравилось - отображение таблиц через транспонирование, настройка моделей -

<https://www.kaggle.com/dejavu23/house-prices-plotly-pipelines-and-ensembles>

[Update] 1st place solution with code | Kaggle -

<https://www.kaggle.com/c/imaterialist-fashion-2019-FGVC6/discussion/95247#latest-549786>

GitHub - amirassov/kaggle-imaterialist: The First Place Solution of Kaggle iMaterialist (Fashion) 2019 at FGVC6 - <https://github.com/amirassov/kaggle-imaterialist>

The problem with metrics is a big problem for AI · fast.ai -

<https://www.fast.ai/2019/09/24/metrics/>

Megafon concurs - <http://bigdatacamp.megafon.ru/>

50 оттенков matplotlib — The Master Plots (с полным кодом на Python) / Хабр -

<https://m.habr.com/ru/post/468295/>

Datafest novosib 2018 — посмотреть видео - <https://datafest.ru/siberia/>

Attention Maps.ipynb — Colaboratory -

[https://colab.research.google.com/github/zaidalyafeai/AttentionNN/blob/master/Attention\\_Maps.ipynb#scrollTo=WC4Ts9mSsKQI](https://colab.research.google.com/github/zaidalyafeai/AttentionNN/blob/master/Attention_Maps.ipynb#scrollTo=WC4Ts9mSsKQI)

Какие закономерности находят нейросети? / Хабр - <https://m.habr.com/ru/post/467895/>

G5 iot - <https://habr.com/ru/post/467317/>

<https://informika.ru/files/contentfile/671/tekst-posobiya-versiya-5.pdf>

Fwd: Prepare data for ml - <https://morioh.com/p/6e560ad84676/how-to-prepare-your-dataset-for-machine-learning-in-python>

Словарь слов - обработка слов и ошибок - <https://proglb.io/p/python-dictionary-app/>

Подготовка данных и дисперсия - <https://prglb.ru/22bk2>

Nlp practice - <https://prglb.ru/5pjkw>

Fwd: Linux command list - [https://vk.com/photo-](https://vk.com/photo-54530371_457252495?api_access_key=0b2bb747ff3ed8718e)

[54530371\\_457252495?api\\_access\\_key=0b2bb747ff3ed8718e](https://vk.com/photo-54530371_457252495?api_access_key=0b2bb747ff3ed8718e)

[https://psv4.userapi.com/c848216/u16487904/docs/d18/eb10838f4df8/Linux\\_komandy.pdf?extra=0P0w616hKzN09SfRcXBn0cZvQL1rw2SDWSowg\\_NGUxLlovalaHfsvwv\\_OpQ\\_paiCvAw0fwq2UQ\\_3lHZGAoHHpRRHq7zGSU\\_DfQhIcYvJAUndPAXDDZnu04RgWH9P1Rlxolq0vPanghgcWw](https://psv4.userapi.com/c848216/u16487904/docs/d18/eb10838f4df8/Linux_komandy.pdf?extra=0P0w616hKzN09SfRcXBn0cZvQL1rw2SDWSowg_NGUxLlovalaHfsvwv_OpQ_paiCvAw0fwq2UQ_3lHZGAoHHpRRHq7zGSU_DfQhIcYvJAUndPAXDDZnu04RgWH9P1Rlxolq0vPanghgcWw)

[https://www.experiment-fest.ru/ab\\_course](https://www.experiment-fest.ru/ab_course) - [https://www.experiment-fest.ru/ab\\_course](https://www.experiment-fest.ru/ab_course)

## SQLITE links

- <https://habr.com/ru/post/149356/>
- <https://python-scripts.com/sqlite>
- <https://proglib.io/p/sqlite-tutorial/>
- [https://wiki.dieg.info/python\\_i\\_sqlite](https://wiki.dieg.info/python_i_sqlite)
- <https://blog.thedataincubator.com/2018/05/sqlite-vs-pandas-performance-benchmarks/>
- <https://habr.com/en/post/321510/>
- <https://pythonspot.com/sqlite-database-with-pandas/>
- <https://pythonworld.ru/obrabotka-dannyx/pandas-cookbook-8.html>
- <http://trainingweb.ru/page/create-database-sqlite-python>
- <https://plot.ly/ipython-notebooks/big-data-analytics-with-pandas-and-sqlite/>
- <https://younglinux.info/python/feature/sqlite>
- <https://switch-case.ru/66640722>
- <https://qaseta.ru/q/0/25902/%D0%BA%D0%B0%D0%BA-%D1%81%D0%BE%D1%85%D1%80%D0%B0%D0%BD%D0%B8%D1%82%D1%8C-pandas-dataframe-%D0%B2-sqlite-db>
- <http://datareview.info/article/otbor-priznakov-dlya-mashinnogo-obucheniya-na-python/>
- <http://python.su/forum/topic/22630/>
- <https://ruseller.com/lessons.php?id=2277>

\*\*\*\*\*

## ML links

Numpy cool article - <https://morioh.com/p/0495f3f41cae>  
Puthon and ml - <https://morioh.com/p/80f9648881bf>  
Open cv full book - <https://morioh.com/p/bc60494297f6>  
Face detection on ml - <https://morioh.com/p/23fe926c7134>  
Top10 algoritms for ml - <https://morioh.com/p/1013a5f310dd>

Установленные пакеты:

- Guake terminal
- Subline Text

```
docker run -it --rm -e JUPYTER_ENABLE_LAB=1 --name jplab --workdir /home/jovyan/work -v  
"$PWD":/home/jovyan/work -p 8888:8888 jupyter/datascience-notebook:9b06df75e445 start-notebook.sh --  
NotebookApp.token="" --NotebookApp.password="" --NotebookApp.allow_origin=* --NotebookApp.base_url=/jp
```

описание наборов юпитер + пакеты

<https://jupyter-docker-stacks.readthedocs.io/en/latest/using/selecting.html>

jupyter notebook - commands

<https://proglib.io/p/jupyter/>

config jupyter for external usefull Доступ к Jupyter notebook через Docker и Nginx. -

<https://abogomolov.com/post/171877094584/jupyter-via-docker-nginx>

videos about jupyter -

<https://hub-binder.mybinder.ovh/user/jupyterlab-jupyterlab-demo-ub1arrze/lab>

Git учебник

<https://git-scm.com/book/ru/v1/%D0%9E%D1%81%D0%BD%D0%BE%D0%B2%D1%8B-Git-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5-Git-%D1%80%D0%B5%D0%BF%D0%BE%D0%B7%D0%B8%D1%82%D0%BE%D1%80%D0%B8%D1%8F>

Телеграм бот - Телега для датасайен

<https://m.habr.com/ru/company/ods/blog/462141/>

Как мы обучили нейронную сеть классифицировать шурупы / Хабр - <https://m.habr.com/ru/company/oleg-bunin/blog/455650/>

Спасательный круг для собеседования на Python-разработчика

<https://proglib.io/p/py-interview/>

Списки команд для Люникса

[https://psv4.userapi.com/c848216/u16487904/docs/d18/eb10838f4df8/Linux\\_komandy.pdf?extra=0P0w616hKzN09SfRcXBn0cZvQL1rw2SDWSowg\\_NGUxLllovalaHfsvwv\\_OpQ\\_paiCvAw0fwq2UQ\\_3lHZGAoHHpRRHq7zGSU\\_DfQhIcYvJAUndPAXDDZnu04RgWH9P1Rlxolq0vPanghgcWw](https://psv4.userapi.com/c848216/u16487904/docs/d18/eb10838f4df8/Linux_komandy.pdf?extra=0P0w616hKzN09SfRcXBn0cZvQL1rw2SDWSowg_NGUxLllovalaHfsvwv_OpQ_paiCvAw0fwq2UQ_3lHZGAoHHpRRHq7zGSU_DfQhIcYvJAUndPAXDDZnu04RgWH9P1Rlxolq0vPanghgcWw)

[https://vk.com/photo-54530371\\_457252495?api\\_access\\_key=0b2bb747ff3ed8718e](https://vk.com/photo-54530371_457252495?api_access_key=0b2bb747ff3ed8718e)

How to Implement CycleGAN Models From Scratch With Keras

<https://machinelearningmastery.com/how-to-develop-cycle-gan-models-from-scratch-with-keras/>

## Видео участия в соревнованиях

Kaggle Santander Customer Transaction Prediction — Василий Рязанов

[https://www.youtube.com/watch?v=yHqzQEwxm3s&feature=share&fbclid=IwAR3uZa3NOvHxwmy\\_VEIUI-Y7at4n5VOVuv2ahakSOKaxlyf4TT6wlruTp4](https://www.youtube.com/watch?v=yHqzQEwxm3s&feature=share&fbclid=IwAR3uZa3NOvHxwmy_VEIUI-Y7at4n5VOVuv2ahakSOKaxlyf4TT6wlruTp4)

Дайджест соревнований по анализу данных – Сергей Брянский - transactions

<https://www.youtube.com/watch?v=2qUkf1VuzHE&feature=youtu.be&fbclid=IwAR2IHVs8YEofUOcKCf8QUfPso1l1sCAnoS6RpxKiQOG3YZx2wDzByHLeAZA>

Kaggle Allstate Claims Severity: предсказание тяжести страхового случая — Алексей Носков

<https://www.youtube.com/watch?v=p7ArDjMImil>

**Kaggle Home Credit: определение риска дефолта по кредитам — Евгений Патеха**

<https://www.youtube.com/watch?v=H4sl3iMTCns>

**Tinkoff Data Science Challenge: анализ задачи предсказания выбора кредита — Алексей Чернобровов**

<https://www.youtube.com/watch?v=1sSr6Bs8rZo>

**ML Boot Camp III: предсказание ухода пользователей из онлайн-игры — Михаил Карачун**

<https://www.youtube.com/watch?v=c7tpExgE2gY>

**Kaggle Two Sigma Financial Modeling Challenge: предсказание финансовых величин — С. Семёнов**

<https://www.youtube.com/watch?v=tZhkw40gDkE>

**Kaggle Corporación Favorita Grocery: предсказание продаж сети супермаркетов — Андрей Филимонов**

[https://www.youtube.com/watch?v=\\_C\\_vXZ\\_y6W8](https://www.youtube.com/watch?v=_C_vXZ_y6W8)

**Определение дома и работы клиента по банковским транзакциям – Сергей Гайдаенко**

<https://www.youtube.com/watch?v=yxOgtGzoTdk>

**OpenAI Retro: как казаки retro контекст решали — Сергей Колесников**

<https://www.youtube.com/watch?v=N5ewwq5TCGE>

**Kaggle Santander: предсказание размера транзакций потенциальных клиентов — Анатолий Ильенков**

<https://www.youtube.com/watch?v=LNDr2GO3PdY>

**Kaggle Mercedes Benz: предсказание времени тестирования автомобилей — Данила Савенков**

<https://www.youtube.com/watch?v=HT3QpRp2ewA>

**Чернобровов Подход к автоматизации решений задач ML на примере HackerEarth ML**

<https://youtu.be/Z3XmVrIHe3c>