

Лабораторная работа №3

Данные (только для задания 3): папка и csv-файл на <https://disk.yandex.ru/d/8XV2OqX2k57lKg>

Дедлайн: 9 июня 23.59 АоЕ (10 июня в 15.00 по МСК)

Форма для сдачи заданий 2 и 3: <https://forms.yandex.ru/u/66511055068ff082b30b70ec/> (после закрытия присылайте в тг нашему ассистенту: @letit66)

Внимание! Эта лабораторная состоит из трех частей. Для получения зачета, как и всегда в Вышке, достаточно получить 4 балла за любое задание или любую комбинацию заданий. Если вы получите больше 10 баллов, дополнительные баллы также засчитаются в состав итоговой оценки.

Задание 1 (5 баллов)

Квиз: <https://forms.gle/5CegAZNZnh9a6R8M6>

Для этого квиза действуют обычные правила сдачи: засчитывается вторая попытка, если только первая не была лучше.

Задание 2 (5 баллов)

Выберите статью, отвечающую вашим научным интересам, и референтируйте её. В идеале она должна быть связана с вашим будущим проектом. В статье обязательно должны применяться методы машинного обучения, но исследование не обязательно должно фокусироваться именно на них.

Возможные источники статей:

- 1) ACL Anthology. Кроме больших конференций, можете обратить внимание на группы специальных интересов - special interest groups (SIGs). Например, есть SIG по цифровым гуманитарным наукам: <https://aclanthology.org/sigs/sighum/>.
- 2) <https://paperswithcode.com/>
- 3) Материалы конференции “Диалог”: <https://www.dialog-21.ru/>
- 4) Любые другие статьи, найденные вами в вашем любимом поисковике;
- 5) Со мной можно согласовать реферирование научно-популярных материалов, например, статей в исследовательских блогах или даже видеолекций. Примером такого материала может служить, например, вот эта статья: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>.

Объем реферата: минимум 2500, максимум 5000 знаков с пробелами. Не забудьте указать статью, которую референтируете. Язык реферата: русский или английский.

Задание 3 (максимум 10 баллов)

Задание на программирование: <https://colab.research.google.com/drive/1nGT30ea4xU6TMqtfGUNIwj35D2GBx8pM?usp=sharing>

Внимание! В задании про конструирование нейросетей уже написана значительная часть вспомогательного кода. Вы не обязаны сохранять этот код. Можете переписать всю тетрадку с нуля, если хотите. Главное – при-
держиваться задачи (бинарная классификация тональности по заданному датасету) и реализовать требуемые архитектуры нейросетей.