

Эссе и проект по курсу “Машинное обучение”

Эссе – исследовательская обзорная работа по одной (или больше, в случае необходимости или особого интереса) статьям. Результатом работы должен быть текст объемом 1-2 страницы.

Проект – практическая работа: вам нужно будет придумать (или взять из собственной курсовой или диплома) задачу, в которой нужно использовать методы машинного обучения и показать, как именно и какие методы следует использовать для решения этой задачи. Кроме того, вам нужно будет продемонстрировать, какие результаты получаются при использовании этих методов и в каких случаях возникают ошибки.

Эссе и проект можно выполнять в группах. Оформление эссе и презентаций по проектам – свободное, но без излишеств. Стоит учитывать, что обе работы носят формальный характер и должны удовлетворять минимальным стилистическим требованиям. Плагиат (любое недобросовестное заимствование текста, кода или изображений без ссылки на источник) не допускается и приводит к обнулению оценки.

1 Эссе по статье

Рекомендуемая структура эссе:

1. Описание лингвистической задачи, о которой идет речь в статье (машинный перевод, распознавание речи, разрешение омонимии);
2. Неформальная постановка задачи (что нужно сделать, какую модель построить или обучить, на каких данных). Оригинальная ли такая постановка задачи или она появляется в более ранних работах?
3. Формальная постановка задачи (оценить вероятность того-то, найти такое-то разложение такой-то матрицы, оценить параметры такой-то модели);
4. Метод решения сформулированной задачи;
5. Эксперименты и результаты;
6. Ваше впечатление: насколько понятна статья, насколько воспроизводимы эксперименты, доступны ли данные, есть код в открытом доступе;
7. Как вы оцениваете значимость статьи для лингвистики и ее оригинальность? Носит ли статья прикладной или более фундаментальный и теоретический характер? Если статья более прикладная, то на каких теоретических работах она основана? Если, наоборот, статья носит более фундаментальный характер, в каких последующих работах использованы предложенные теоретические положения?

8. Если код в открытом доступе, пример запуска кода, если его можно запустить вообще и ваша оценка, насколько хорошо метод справляется с поставленной задачей
9. Ваши предложения: какие эксперименты еще можно было бы провести, какие слабые места вы видите в работе, как можно усовершенствовать предложенную модель, где бы вы могли использовать описанный в статье метод. Если эксперименты посвящены языку, отличному от русского, насколько трудно было бы перенести полученные результаты на русскоязычные данные?

2 Проект

Рекомендуемая структура презентации по проекту: Презентацию можно оформить в виде jupyter notebook со всеми необходимым комментариями или презентации.

1. Постановка задачи и описание используемых данных;
2. Описание использованной внешней библиотеки или разработанного решения;
3. Демонстрация решения (на одном слове, предложении, небольшом тексте);
4. Оценка результатов в целом: формальные показатели качества (например, ассигасу) и неформальная оценка результатов: насколько хорошо или плохо получилось решить поставленную задачу;
5. Анализ ошибок: в каких случаях возникают ошибки и почему?
6. Ваши ощущения: какие возникли сложности, удалось ли с ними справиться с поставленной задачей, сколько времени ушло у вас на выполнения проекта, можете ли вы порекомендовать использованное внешнее решение или самостоятельно разработанное решение коллегам;
7. Если все хорошо и все получилось, то что можно делать с готовым решением и где и кто его будут использовать, а если совсем все плохо и ничего не получилось, то почему и как с этим жить дальше.