

Оригинальные решения при проведении контрольных/оценочных мероприятий

Ф.И.О. (полностью) преподавателя, подающего заявку

Романов Александр Юрьевич

Кампус/факультет/школа/департамент, должность

Национальный исследовательский университет «Высшая школа экономики» / Московский институт электроники и математики им. А.Н. Тихонова / Департамент компьютерной инженерии, Доцент ДКИ МИЭМ НИУ ВШЭ

Укажите контактный телефон, по которому с вами можно связаться

+79161961722

Укажите e-mail, по которому с вами можно связаться

a.romanov@hse.ru

ФИО, должность соавтора(ов), если имеются

Зунин Владимир Викторович, ассистент ДКИ МИЭМ НИУ ВШЭ

Название учебной дисциплины и название образовательной программы, где читается курс

Системы искусственного интеллекта, ОП Информатика и вычислительная техника

Программа учебной дисциплины

<https://www.hse.ru/pubs/share/folder/c76e555abcedf5/837036602>

Средний балл итоговых оценок студентов по курсу, в рамках которого применялась описываемая практика

6,37 (по 10-тибальной системе) / 4,0 (по 5-тибальной системе)

Заявка подается впервые или после доработки?

впервые

Если заявка подается после доработки, то опишите подробно произведенные изменения (добавление новых элементов, изменения в материалах и т.п.)

-

Название описываемой практики

Использование платформы Kaggle для организации конкурсов по исследованию данных для проведения и оценивания практических работ по дисциплине «Системы искусственного интеллекта».

Какие цифровые инструменты используются в данном курсе?

Kaggle – это ведущая платформа для организации конкурсов по исследованию данных. В настоящее время принадлежит корпорации Google. На этой платформе преподаватель создает закрытое соревнование, где размещает задание на практическую работу, справочные материалы и датасет. Студенты ведут разработку моделей машинного обучения на языке Python и отправляют свое решение на платформу. Эти решения автоматически тестируются на скрытом валидационном наборе данных и рассчитываются метрики эффективности присланных решений. Kaggle бесплатно предоставляет ресурсы времени работы мощных вычислителей на основе GPU (в основном недоступных на обычных компьютерах), на которых студенты могут обучить и выполнить инференс своих моделей. Также Kaggle предоставляет специальные инструменты и библиотеки для исследования данных и машинного обучения. Происходит автоматическое рейтингование решений и строится рейтинговая таблица, на основе которой в систему ЛМС, по завершению соревнования, студенту выставляется оценка. Студент может видеть свой рейтинг и рейтинг других участников, дорабатывать решение и отправлять улучшенную версию, соревнуясь с одноклассниками. Студенты ограничены только сроками выполнения каждой практической работы, определяемыми преподавателем. Присланные решения, исходные коды предложенных решений, могут дополнительно проверяться преподавателем на плагиат. У каждого соревнования Kaggle есть свой отдельный чат/форум, где студенты могут обсудить задание, задать вопросы и получить консультацию.

Лабораторные работы подбирались в соответствии с программой дисциплины, которая охватывает базовые технологии и направления по теме: «Системы искусственного интеллекта». Было четыре основные работы, разделенные по данным, которые обрабатывались:

- 1) Табличные данные. Необходимо было решить задачу предсказания вероятности оттока клиентов организации по табличным данным.
- 2) Работа с изображениями. На вход поступали изображения, на которых могли находиться люди. Необходимо было разработать детектор человека на изображении.
- 3) Работа с текстом. Был дан набор рецензий к фильмам. Необходимо было разделить полученные рецензии на два класса: положительные и отрицательные.
- 4) Работа с аудио. Студентам был предложен набор аудиозаписей, каждая из которых содержит одно из десяти слов. Студенты должны были разработать решение для определения произносимого слова.

Для каждой из представленных работ разработаны свои критерии для оценки студентов. Общая схема оценивания выглядит следующим образом:

- 1) Для оценки «удовлетворительно» студентам достаточно было использовать решения, которые разбирались на занятии, нужно было реализовать их в виде программы.
- 2) Для оценки «хорошо» требовалось изучить дополнительные материалы, выданные в виде списка литературы для дополнительного изучения.
- 3) Для оценки «отлично» студентам необходимо было провести самостоятельный поиск возможных решений и реализовать их в виде программы.

Например, для первой работы, где надо было предсказывать вероятности оттока клиентов организации, в качестве базового решения предлагалось использовать функции из библиотеки XGBoost или нейронную сеть. Для оценки «удовлетворительно» достаточно было разобраться, как работает библиотека XGBoost и разработать программу с ее использованием. Для оценки «хорошо» требовалось придумать, как улучшить базовое решение. Для оценки «отлично» нужно было использовать более продвинутую модель машинного обучения, или комплекс моделей.

У преподавателя есть доступ к решениям студентов, а также автоматизированные средства для дополнительной проверки при подозрении на плагиат. Это исключает возможность списывания.

В качестве тестовой выборки для тестирования решений студентов были использованы данные из открытых источников. Полученные данные были сформированы в единый датасет и модифицированы различными способами: изменения названия полей, добавления каких-либо специфических данных или корректировка имеющихся и т.д. Поэтому тестовую выборку можно считать уникальной, и она доступна только преподавателю. Это позволяет говорить о том, что утечка данных невозможна, и обмануть тестовую систему нельзя, за счет чего достигается справедливое оценивание студентов.

На достижение каких образовательных результатов нацелена описываемая практика/нововведение?

Студенты учатся пользоваться современной общепризнанной международной платформой для проведения соревнований и исследований по Искусственному интеллекту. Студенты овладевают навыками поддержки программных проектов, ведения Github-репозитория, соблюдения Code style, и другими практиками профессиональной IT-разработки.

По каким основаниям вы полагаете, что предложенная методика приводит к планируемому результату обучения?

Внедрена автоматизация проведения практических работ по дисциплине «Системы искусственного интеллекта», что снижает нагрузку на преподавателя. Это позволило обеспечить автоматическое оценивание и объективное человеко-независимое рейтинговое присланных решений (схема и процедура оценивания прозрачны, понятны и одобряемы студентами). Выполнение практических работ не привязано к оборудованию лаборатории и времени проведения аудиторных занятий. Это позволяет преподавателю в аудитории уделить больше времени консультированию студентов и снизить нагрузку на лабораторный фонд.

По каким основаниям вы можете оценить, что описываемая практика оказалась эффективной по итогам реализации курса?

Дисциплина «Системы искусственного интеллекта» преподается в 1-2 модуле 4 курса ОП «Информатика и вычислительная техника» специализации «Автоматизированные системы» (акцент на разработке строящихся систем и проектировании СБИС/ПЛИС). Дисциплина организована таким образом, чтобы дать будущим разработчикам электронных средств основы машинного обучения, но ее наличие и содержание настолько влияют на студентов, что они в качестве ВКР охотно выбирают темы связанные с машинным обучением. Например, темы ВКР 2023 г.: «Оценка надежности комбинационных схем с использованием методов машинного обучения», «Разработка датасета реализаций процессорной архитектуры RISC-V», «Разработка нейросетевого алгоритма для синтеза и разметки речи», «Анализ изображений ТРУЗИ простаты с использованием методов машинного обучения»,

«Автоматическое формирование ключевых слов для русскоязычной научной литературы», и др. До этого курс преподавался другим преподавателем, и имел низкие оценки/отзывы по причине устаревшей программы и преподаваемых технологий в нем. В этом году он был передан другим преподавателям. Курс был полностью переработан и приведен в соответствие с современными требованиями и технологиями. Не смотря на то, что дисциплина была новой и довольно сложной для освоения студентами (по скольку является дополняющей для их профиля), она была встречена с большим интересом, а методика проведения практических работ получила одобрение и положительный отклик от студентов. Это отразилось на оценке СОП по этой дисциплине, которая из низкой (до изменения преподавателя и методики преподавания), выровнялась до среднего значения по департаменту за учебный период. Оценки студентов распределились близко к графику функции Гаусса (средний балл: 6,37), что свидетельствует о сбалансированности системы оценивания и отсутствии проблемы «инфляции оценок».

Средний балл СОП по курсу, в рамках которого применялась описываемая методика

3,76

Прочие файлы\документы по Вашей заявке

<https://www.hse.ru/pubs/share/folder/7987c63bf23fda/837036604>