**Проектное предложение**

|  |  |
| --- | --- |
| Тип проекта | Исследовательский |
| Название проекта | **Поиск коррупционного контента в русскоязычных и англоязычных СМИ с помощью методов искусственного интеллекта** |
| Подразделение инициатор проекта | Проектно-учебная лаборатория антикоррупционной политики |
| Руководитель проекта | Доктор социологических наук, кандидат психологических наук, Александр Александрович Максименко, аналитик Проектно-учебной лаборатории антикоррупционной политики ВШЭ |
| Заказчик проекта / востребованность проекта | Проектно-учебная лаборатория антикоррупционной политики (ПУЛАП) НИУ ВШЭ |
| Основная проектная идея / описание решаемой проблемы | Основная проектная идея состоит в ознакомлении с алгоритмами работы искусственного интеллекта по автоматическому поиску новостного контента по теме коррупции в русскоязычных и англоязычных СМИ. В ходе проекта решается проблема классификации новостного контента по теме коррупции и разметка собранного медийного массива, удобного для машинного обучения. |
| Цель проекта | Цель проекта: попытка создания модели с использованием искусственного интеллекта для поиска и классификации новостного контента, содержащего упоминания коррупции в русскоязычных и англоязычных СМИ.Задачи проекта:1. Познакомить студентов с особенностями семантического (в том числе полимодального) анализа;2. Познакомить студентов с методами искусственного интеллекта;3. Создать автоматический новостной поиск по теме противодействия коррупции в русскоязычных и англоязычных СМИ с помощью методик искусственного интеллекта. |
| Планируемые результаты проекта, специальные или функциональные требования к результату | Размеченный новостной массив, собранный в русскоязычных и англоязычных СМИ (не менее 300 новостей) и обратная связь о работе алгоритмов ИИ (искусственного интеллекта). |
| Требования к участникам с указанием ролей в проектной команде при групповых проектах | Знание английского языка (базовый уровень), желательное владение методами контент-анализа, а также хорошая ориентация в новостных русскоязычных и англоязычных сайтах. |
| Количество вакантных мест на проекте | 25 |
| Проектное задание | Проект предусматривает поиск не менее 300 новостей по коррупции в русскоязычных (150) и англоязычных (150) СМИ по заданным параметрам и ключевым словам за последние 5 лет, их разметку (классификацию). С помощью размеченного массива происходит обучение искусственного интеллекта осуществлять автоматический поиск новостей в русскоязычных и англоязычных Интернет-СМИ и соотносить обнаруженные новости с принятой классификацией. |
| Критерии отбора студентов | Заинтересованность в теме проекта, научная любознательность и исследовательский интерес |
| Сроки и график реализации проекта | 01.10.2020 – 01.02.2021 |
| Трудоемкость (часы в неделю) на одного участника | 8 часов в неделю |
| Количество кредитов | 4 кредита |
| Форма итогового контроля | Экзамен (выставление оценок в ведомости) |
| Формат представления результатов, который подлежит оцениванию | Массив в формате xls с указанием ссылок на новостные порталы (150 русскоязычных + 150 англоязычных) с новостями, содержащими упоминание о коррупции за последние 5 лет; отчет (в формате doc) о работе алгоритмов ИИ в ходе автоматического определения новостного контента нацеленного на поиск с указанием возможностей и ограничений. |
| Виды деятельности, выполняемые студентом в проекте/отрабатываемые навыки | Умение самостоятельно найти новостные источники по теме проекта на русском и английском языках, проанализировать новостной контент, разметить (классифицировать) собранный массив, проанализировать работу методик искусственного интеллекта по поиску новостного контента с указанием возможностей и ограничений методики с учетом перспективных исследований. |
| Образовательные результаты проекта | Получение навыков собирать и анализировать данные СМИСпособность анализировать значительные массивы новостного контента, выявляя по заданным критериям существенные для классификации данные и разметки тексты Способность выявлять типичные коррупционные практики в новостном контентеПонимание новостей о коррупции на английском языке |
| Критерии оценивания результатов проекта с указанием всех требований и параметров | Количество новостных источников на русском и английском языках по теме коррупции – 0,7; грамотная (адекватная) разметка (классификация) собранного массив по теме коррупции – 0,2; анализ работы методики искусственного интеллекта на основе собранного массива – 0,1. |
| Возможность пересдач при получении неудовлетворительной оценки | Да |
| Рекомендуемые образовательные программы | Для любых образовательных программ, предпочтение студентам ОП:41.00.00 «Политические науки и регионоведение», 37.00.00 «Психологические науки»,39.00.00 «Социология и социальная работа», 40.00.00 «Юриспруденция», 38.00.00 «Экономика и управление»,42.03.02 «Журналистика». |
| Территория | РФ, г. Москва, Мясницкая ул., ауд. 407 |