***ФГБОУ ВО «Кубанский государственный аграрный университет имени И.Т. Трубилина»,***

***Российская Федерация***

**198 Практическое занятие №5. По дисциплине "Интеллектуальные информационные системы и технологии". 2020-11-11**

**Заголовок:**  
Обсуждение проекта по анализу текстов детских книг (Kaggle) для курсовой и публикаций во время онлайн-занятия

**Резюме текста:**

**1. Начало занятия и отсутствие студентов**  
Преподаватель начинает онлайн-занятие по дисциплине "Интеллектуальные системы и технологии" для группы ПИ-19-41 11 ноября 2020 года. Отмечает отсутствие студентов и предполагает, что занятие может быть сорвано. Выясняется (после подключения одной студентки, Насти), что группа отсутствует из-за проходящего в этот день тестирования в рамках аттестации/аккредитации. Преподаватель соглашается, что проведение занятия в такой день неуместно.

**2. Обсуждение проекта на основе данных Kaggle**  
Преподаватель делится идеей для курсовой работы или проекта, основанной на данных с платформы Kaggle.

* **Предложение темы и данных:** Предлагается использовать датасет с аннотациями к детским книгам.
* **Суть задачи и успешность моделирования:** Задача – автоматически определять возрастной диапазон читателя по тексту аннотации книги. Преподаватель сообщает, что уже построил успешную и достоверную модель для решения этой задачи прямо во время предыдущего занятия.
* **Технические детали:** Модель использует слова из аннотаций как признаки для классификации. Обсуждаются настройки ввода данных в системе Eidos, в частности, необходимость указать интерпретацию текстовых полей и использование слов как признаков.

**3. Применение проекта и публикация результатов**  
Обсуждается практическое применение предложенного проекта.

* **Использование для курсовой работы:** Проект отлично подходит для курсовой работы, так как задача интересная, данные доступны, а решение реализуемо и наглядно.
* **Возможности публикации:** Преподаватель настоятельно рекомендует публиковать результаты исследований, упоминая платформы ResearchGate и РИНЦ. Он делится ссылкой на обновленную инструкцию и шаблон для оформления научных статей, подчеркивая простоту процесса публикации.
* **Связь с требованиями аттестации:** Подчеркивается актуальность навыков представления научных результатов (ОПК-5), которые проверяются во время текущей аттестации магистрантов и аспирантов, и как предложенный проект помогает их развить.

**4. Рефлексия преподавателя**  
Преподаватель размышляет о том, что часто предвидит важные направления развития (как, например, необходимость публикаций и представления исследований), но сталкивается с медленным внедрением этих идей в академической среде, что вызывает у него грусть и сожаление об упущенных возможностях.

**5. Обсуждение инструментов и завершение**

* **Конвертер CSV в XLS:** Преподаватель делится ссылкой на удобный онлайн-конвертер для преобразования файлов CSV (формат данных с Kaggle) в XLS, отмечая его преимущества перед другими инструментами и встроенными возможностями Eidos при работе с текстами, содержащими запятые.
* **Прощание и пожелания:** Преподаватель и студентка прощаются. Преподаватель желает удачи на тестировании и завершает запись занятия, констатируя отсутствие остальных студентов.

**Детальная расшифровка текста:**

**1. Начало занятия и отсутствие студентов**

Так.  
Уже запись идёт на самом деле.

Так, ребята.  
Сегодня 11 ноября 2020 года. Вторая пара, 9:45-11:15. Практическое занятие номер пять с группой ПИ-19-41. По дисциплине Интеллектуальные системы и технологии.

Никого нет присутствующих.

Ждём 15 минут до 10 часов. Если никого не будет, занятие считается сорванным.

Ну сейчас идёт аттестация, аккредитация, тестирование. Возможно, они на тестировании. Просто можно было бы предупредить, конечно. Ну, как есть.

Ну это если они не придут за 15 минут. Ну я думаю, что они не придут. Они вообще ходят так себе. Ни шатко, ни валко.

*[Звук уходящего человека]*  
Ну пошли.  
Ну аккуратно, пожалуйста.

*[Звук подключения нового участника]*  
Настя, привет.  
Здравствуйте.  
Ну что у нас там? Где-то группа?  
Не знаю, подключится кто-то или не подключится. У нас сегодня тестирование.  
А, в вашей группе даже?  
Да-да-да. Скорее всего, все готовятся.  
Ну понятно. Понятно.

Я тоже так думаю, что сейчас как бы не самое время занятия проводить. Сейчас вот у меня была группа, вот закончилось занятие, ну первая пара. И через 15 минут у них там идти надо уже в аудиторию и садиться уже готовиться к тестированию.  
Ну вот у нас такая же ситуация получается.  
Я думаю, что как бы неуместно, конечно.  
Угу.  
Вот, сейчас это всё как бы второстепенным является. Надо тестироваться хорошо.

**2. Обсуждение проекта на основе данных Kaggle**

**2.1. Предложение темы и данных**

И, знаешь, о чём я хотел сказать? Я сейчас вот с этой группой предыдущей работал и нашёл очень хорошее приложение. Может, нужно? Чтобы кто-нибудь возьмёт его?  
Какое?  
Ну там на Kaggle детские книжки, база данных к детским книжкам, аннотации к детским книжкам.

**2.2. Суть задачи и успешность моделирования**

И нужно понять, вот они, какие особенности, которые описывают книжки каждого возрастного диапазона, можем ли мы раскрыть правила, для которого мы можем автоматически назначить возрастной диапазон для новой книги. Прекрасно получилось. Достоверная модель, всё чётко. То есть просто она слова рассматривает как признаки и связывает с возрастной категорией.  
Угу.  
И выясняется, что оказывается, вполне эта задача разрешимая. То есть я в один момент её там решил, прямо на занятии.

**2.3. Технические детали (Eidos, признаки)**

Сейчас тебе ссылочку кину. Там единственное, что надо данные, в смысле, когда данные вводишь, задать с правой стороны специальную интерпретацию текстовых полей, рассматривать слова как признаки. И всё, понимаешь? А больше ничего делать и не надо будет.

**3. Применение проекта и публикация результатов**

**3.1. Использование для курсовой работы**

Имеется в виду на курсовую, да?  
Ну, на... ну можно на курсовую, можно на это... просто приложение сделать.  
А можете скинуть мне в Ватсапе или где-нибудь на почту?  
Да, сейчас, сейчас тебе пошлю. Вот, короче, и получилось очень интересно. Я сейчас вот только что сделал модель. Они уже занятие закончилось, они ушли.  
Угу.  
Вот. Детские книжки, рассказы с высокими оценками. Там, короче, суть задачи такая: есть книжки, причём много их, там что-то 400 что ли там или... Ну, короче, так сравнительно много. Значит, и они соответствуют различному возрасту, эти книжки. И там текст об этой книжке, аннотация такая небольшая, о чём книжка. Есть ли там картинки, там, что-то такое вот.

Она точно как лабораторная работа. Я даже подумал, может, лабораторную работу заменить этой. Потому что она такая же точно, как лабораторная работа, которую я провожу по...  
Можно её взять в качестве курсовой?  
Да запросто.  
Просто...  
Вот. Очень хорошая задачка, небольшая такая база данных. Она хорошая в обработке, обработка быстро происходит.

**3.2. Возможности публикации (ResearchGate, РИНЦ) и шаблон оформления**

Вот. И я ещё инструкцию модернизировал, и там новый шаблон. Сейчас я тебе пошлю на инструкцию ссылочку, как вот это всё вообще оформлять. Вот. Ну одно другому, кстати, не мешает. Можно вполне и курсовую сделать, и опубликовать её. Почему бы и нет? Потому что ты-то можешь зарегистрироваться в этом ResearchGate. Я надеюсь. Или у тебя тоже нет корпоративной почты? Вот.  
Не, у меня есть корпоративная почта.  
Ну зарегистрируйся в ResearchGate, будешь публиковаться, когда тебе захочется. Вот у тебя возникнет такое желание опубликовать, ты сможешь это сделать.  
Ну мне сейчас в ближайшее время публиковаться будет проблематично.  
ResearchGate размещается за 5 минут всё, понимаешь?  
Не, я понимаю...  
После этого отправляешь... после этого отправляешь его в РИНЦ, оно там тоже размещается. У тебя появляется публикация. Вообще никаких...  
Я понимаю, просто туда же что-то надо загружать.  
Ну вот написать надо. Так вот, написать вот по шаблону, который вот я сейчас тебе послал.  
Угу.  
Это это статья уже написана студентами.

**3.3. Связь с требованиями аттестации (ОПК-5)**

Вот то, что я всё время так и нычу, там: давай, давай, там, пишите. А сейчас ты знаешь, что идёт сейчас тестирование? Меня вчера вообще замучили. Все пишут мне эти, ну и звонят и говорят: "Присылай тест". Ты знаешь, какой тест?  
Какой?  
Для магистрантов и, сейчас тебе даже зачитаю, и аспирантов одно и то же, ОПК-5, которое проверяется способность профессионально излагать результаты своих научных исследований и представлять их в виде научных публикаций и презентаций. Представляешь себе? Это то, чего я заставляю практически вот и учу этому, как это всё сделать: провести исследование, описать его методически правильно и опубликовать. Прямо вот это всё написано, представляешь? И сейчас это будут проверять.

**4. Рефлексия преподавателя**

А я это пытался вот вас на это дело сподвигнуть с самого начала, понимаешь? То есть я вообще-то правильно понимаю всё. Я в этом не раз убеждался. Причём не просто правильно, а на десятки лет вперёд всё понимаю правильно.  
Ну это же хорошо.  
Хорошо, только грустно очень. Наблюдать, как проректора не въезжают, потом когда они умирают, тогда начинают реализовываться то, что я говорил, только не мной уже. Грустно это, понимаешь, видеть всё.  
Ну если так рассуждать, то да.  
Ну вот, и причём очень много возможностей, колоссальное количество возможностей потерянных просто. Просто неиспользованных в связи с идиотизмом, понимаешь?

**5. Обсуждение инструментов и завершение**

**5.1. Конвертер CSV в XLS**

Ладно, давай, иди готовься.  
Вот.  
Всё, хорошо, спасибо.  
Слушай, а это всё...  
Я скопировала эти ссылки.  
Что-что?  
Ссылки скопировала себе.  
Ага, ну отлично.

И для того, чтобы этот CSV файл преобразовать в XLS, вот этот конвертер онлайновский. Вот. Система Eidos она режет там длину поля почему-то, я, значит, в этом конвертере, который я делал, оно преобразует, но режет поле, и смещение происходит. Там где-то есть в самом, в самом поле есть запятые, понимаешь?  
Угу.  
В самом CSV, там же поля. И вот в поле, в поле есть текст, а в тексте запятые. Всё, она начинает, соответственно, считать, что это тоже фрагмент поля какого-то следующего. И получается, ползёт там всё. А вот этот CSV онлайн, он догадывается, как правильно. Он более развитый, он, конечно.

Да, вот это после... последнее - это то, что...  
Да, это конвертер. Их там много, конечно, в интернете, но этот вот... Я просто уже как бы смотрел...  
Ну да, я уже много их видел, и вот пришёл к тому, что вот этот мне лучше. Причём он большие, он большие файлы преобразует корректно, ничего там не конючит, не просит, чтобы там что-то заплатили. Ну так просят, но очень тактично, не раздражающе.

**5.2. Прощание и пожелания**

Ну хорошо, спасибо. Тогда до свидания, я пойду готовиться.  
До свидания, счастливо, удачи в тестировании.  
Угу, спасибо.  
Счастливо.

Всё, ребят, на этом у нас конец занятия. Это я говорю сам себе, потому что никого нету, все ушли на тестирование. Всего самого-самого хорошего. Надеюсь, занятие понравилось. Ещё бы.