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

***Российская Федерация***

**183 Практическое занятие. По дисциплине "Теория информации, данные и знания". 2020-11-05**

Заголовок: Практическое занятие: Анализ студенческих моделей, руководство по публикации в ResearchGate/РИНЦ

Резюме:

1. Введение и Организационные Моменты:

Занятие проводит профессор Луценко Е.В. для группы ПТ-2002 по дисциплине "Теория информации, данные, знания" 5 ноября 2020 года. Обсуждаются организационные вопросы, включая регистрацию присутствующих и подтверждение номера практического занятия (№7).

2. Основная Задача Занятия:

Главная цель – рассмотрение и анализ моделей, разработанных студентами для их курсовых или проектных работ. Проверяется работоспособность и достоверность моделей.

3. Процесс Утверждения Темы и Публикации:

Студенты выбирают тему и создают модель.

Модель анализируется на занятии с использованием специализированного ПО (подразумевается система "АИДОС-X").

Если модель признается достоверной, тема утверждается.

Студенты готовят описание работы по шаблону.

Работа размещается на платформах ResearchGate и РИНЦ (Российский индекс научного цитирования).

4. Публикация на ResearchGate и РИНЦ:

ResearchGate: Требуется регистрация, для которой часто необходим корпоративный email (.edu). Обсуждаются трудности с получением таких email студентами. Варианты: самостоятельная регистрация (если есть email), публикация через соавтора (староста, преподаватель, сотрудник с email), обращение в ЦИТ вуза за корпоративным email.

РИНЦ: Требуется отдельная регистрация и получение SPIN-кода. Процесс может быть долгим. После получения кода необходимо заключить договор с РИНЦ как физическое лицо для размещения непериодических изданий (статей, описаний). Демонстрируется интерфейс РИНЦ для добавления публикаций. Публикации в РИНЦ важны для портфолио и могут учитываться при назначении стипендий.

5. Использование ПО для Анализа Моделей:

Демонстрируется работа с ПО для анализа данных: загрузка данных из Excel, настройка параметров анализа (классификационные и описательные шкалы, обработка пропусков, выбор интервалов).

Обсуждается наличие неактивных пунктов меню ("в разработке") в текущей версии ПО.

Подчеркивается важность корректной подготовки данных (замена пропусков нулями в числовых колонках).

Интерпретируются результаты: оценка достоверности модели, анализ графиков и дендрограмм для выявления структуры данных и сходства/различия объектов.

6. Анализ Студенческих Работ:

Пример 1 (Александр Перков): Анализ характеристик оружия в компьютерной игре. Данные скорректированы, модель построена и признана достоверной. Тема утверждена.

Пример 2 (Глеб Клёнс): Анализ цен на квартиры в Москве. Данные проанализированы, модель признана достоверной. Тема утверждена. Отмечается, что похожая работа (возможно, этого же студента) уже разбиралась ранее.

Пример 3 (Марина): Анализ доступности контента в разных странах (данные Netflix). Запущен анализ, но из-за большого объема данных (2 Мб текста, много признаков-слов, 500+ классов-стран/регионов) расчет требует значительного времени (оценка >20 минут). Студентке будут высланы параметры для самостоятельного запуска расчета на ее компьютере.

7. Завершение:

Занятие завершается. Студентам, чьи темы утверждены, дается задание готовить описания для публикации. Обсуждение работы Марины отложено из-за длительности расчетов.

Детальная Расшифровка Текста:

1. Введение и Организационные Моменты

Здравствуйте, ребята.

Здравствуйте.

Здравствуйте, здравствуйте.

Здравствуйте.

Сегодня 5 ноября 2020 года, пятая пара с группой ПТ-2002.

Практическое занятие по дисциплине Теория информации, данные, знания.

Занятие ведёт профессор Луценко Евгений Вениаминович.

Сейчас, ребята, немножко я сделаю паузу, потому что я сейчас буду везде там регистрировать, что я присутствую на занятии, что вы присутствуете на занятии. Вот это нужно всё сейчас зарегистрировать это.

Это у нас практическое занятие номер семь.

Пятая пара.

Наверное, у меня сегодня так и есть пятая. Тут первые три, четвёртая, пятая. Замечательно.

Здесь я уже отметился, что я присутствую на занятии.

И система не может... О, смогла.

Так, это у нас, значит, седьмое занятие.

2. Основная Задача Занятия

И у нас сейчас тема всех занятий, которые проходят, это мы смотрим на те модели, которые вы выбрали. Вернее, на те темы, которые вы выбрали, и пробуем создать модель и посмотреть, насколько она нормально работает.

3. Процесс Утверждения Темы и Публикации

Если она получается нормальная, а это, как правило, так и получается, то тогда я утверждаю тему, вы начинаете писать описание по шаблону этого приложения и размещать его в ResearchGate и РИНЦ.

4. Публикация на ResearchGate и РИНЦ

Общие принципы и опции

Вот, и вам такую информацию сообщаю, что у вас есть выбор. Значит, если вы не зарегистрируетесь сами в этом ResearchGate. Если вы сами зарегистрируетесь, то там вообще понятно, просто сами пишете и размещаете статью в ResearchGate, а потом в РИНЦе. Вот. А если вы сами там не зарегистрируетесь, то вы тогда будете нуждаться в соавторе, который это сделает, который там уже зарегистрирован. Ну, это на мне свет клином не сошёлся в этом плане, но я могу это сделать для кого-то, кто сочтёт нужным. Но вообще-то я могу вам сказать, что у вас у старост есть корпоративные адреса электронной почты. Есть ещё также другие преподаватели, сотрудники кафедры, у которых тоже есть корпоративная электронная почта. Можете с ними в соавторстве статью написать. Да.

Проблемы с корпоративной почтой

Евгений Вениаминович, мне почту не выдали. Ну, я вот староста группы, а почту мне корпоративную не выдали сейчас.

Ещё пока нет, да?

Да, нету. Как мне сказал куратор, нам почту корпоративную выдают...

А какой, какой группы ты староста?

ИТ-2002.

А был Герман, да, Штоль?

Да, да.

А теперь ты, да?

Да. Короче, мне э-э...

Чебыкин?

Сказал куратор то, что...

Так, Чебыкин, э-э как зовут? Андрей, да? Андрей.

Не-не, староста Герман у нас.

Староста-то я. Староста Герман Штоль.

А. Ясно. А тогда я не понял, о чём ты говоришь.

Я говорю то, что нам староста сказал, что корпоративку дают с крайней... И то только старостам. А ещё он сказал, что её могут выдать, если мы идём на дистанционное обучение. Но так как мы сейчас не на дистанционном обучении, и почты у меня нет. Никто не может зарегистрироваться в ResearchGate.

А сейчас у нас какое обучение? Не дистанционное, что ли?

Ну, у нас только с вами дистанционное, а в основном с другими преподавателями мы очно занимаемся.

Ясно.

Варианты соавторства и РИНЦ

Значит, ребята, тогда Герману Штолю очень повезло. Он будет соавтором всех статей, всех студентов своей группы. В общем, выбирайте его или меня, короче. Я совершенно спокойно к этому отношусь. То есть выберете его, слава Богу. Ещё можете кого-нибудь выбрать. В общем, это ваш вопрос. Вот, у вас будут статьи совместные. Вот. Но это всё равно будут ваши статьи. Если там кто-то будет соавтором, то всё равно ваша статья будет. Вот, себя на первой позиции ставьте, а вот кого-то там второго на второй позиции.

Вот. Ну и можно ещё попытаться через куратора обратиться в ЦИТ. Вообще-то, на самом деле, эта почта гугловская. Вы ж там знаете, ЕДУ, что-то Google даёт эти домены. Вот.

А в ЦИТе я спрашивал, кстати, там почтами не занимаются для студентов.

А кто ими, кто их даёт?

Не знаю. Сами появляются.

Вексом не занимаются.

Да нет, ну в ЦИТе там есть э-э этот самый, Христианка, Христич, Христич, который прямо отвечает за этот Webex. Как же там не занимаются? То не в тот отдел там подошли там, обратились. Ну кто-то, наверное, не занимается, кто-то занимается. Там же много отделов, подразделений. Да. Ну так вот. Такова жизнь.

Ну, э-э, ребят, значит, вы особо так не расстраивайтесь, делайте эту работу. Вот. Ну я в прошлые годы, я вам показывал, размещал эти материалы. Ну и в этом тоже размещу. Но, но в РИНЦе вы можете зарегистрироваться, это 100%. Только у вас времени не хватает, наверное, или желания, не знаю, что вы.

Вот скажите, пожалуйста, ребята, кто из вас уже зарегистрировался в ЦИТе, в э-э, Господи, извините, в РИНЦе? Пятая пара сказывается. Дайте мне информацию обратную, обратную связь. Кто зарегистрировался в РИНЦе и получил SPIN-код? Вот один тут мне написал уже в чат. Глеб. Но SPIN-код ещё не получил. Не PIN, а SPIN. Science Personal Index. Вот. PIN-код - это на телефоне. Кода ещё нет. Ну вот они с этим тормозят, вообще-то. Вот они, кстати, знаете, они пишут там 10 дней, что ли, или 7 дней, но фактически может быть и дольше, и месяц может быть. Вот. Значит, как только получаете код, сразу начинаете мгновенно, сразу же с этого же дня, э-э, поскольку они такие тормозные, то надо их там долбить, чтобы они что-то сделали. Ни у кого кода нет. Как только код получаете, сразу же вот сюда заходите на эту ссылочку. На неё можно зайти по-разному. Ну вот я вам просто дал готовую ссылочку, а вообще-то можно зайти через главное меню: авторам. Так, я сейчас, у вас есть эта демонстрация экрана, нет?

Есть.

Есть. Вот смотрите, э-э, главное меню этого eLibrary. Вот, ну уже нет смысла вам её давать ещё раз адрес. Авторам. Вот так вот крутим, крутим. Находим размещение непериодических изданий. Крутим, крутим, находим, смотрите, система заключения договоров. И получаем ту ссылочку, которую я вам послал. И здесь вы заключаете договор на физическое лицо. Договор на размещение непериодических изданий. Вот смотрите, договор на книги, непериодические издания. Они, на самом деле, не только книги, но и статьи. Допустим, из этого же самого ResearchGate. Вот этот номер они присваивают. Я его в тринадцатом году заключил. И у вас вот так вот появится такая вот строчечка. И вы начнёте размещать там свои публикации. Смотрим, отображаем, какие там у меня издания есть, которые я там размещал. А там их вот смотрите, ребята, вот я размещал, видите, сколько? Вот это размещал, вот это размещал. Вы поняли уже, да? Вот это размещал. Ну и ещё размещу. И вот возьмём вот эту последнюю. То есть я это умею делать. Значит, если у кого-то будут появляться вопросы, ребят, как это делается, обращайтесь, я буду объяснять. Там вопросов особых нет. Единственный вопрос, когда есть соавторы, то могут возникать вопросы. И ещё такой момент. Вы, значит, можете просто э-э вот здесь вот написать, что вот у меня есть вопрос к сотрудникам. Выбрать тему сообщения, э-э размещение непериодических изданий. Написать: "Я вот там то-то, то-то сделал, а ничего нету. Что мне делать?" И вам скажут: "Зайдите туда-то, нажмите то-то, там, и у вас появится это вот". То есть там суть-то какая? Мы даём им материалы, они на них смотрят, на эти материалы, и принимают решение. Вот, сейчас я вам покажу, как наша, ваше издание. Вот, размещение рецензий ваших, так. Ваше издание. Вот. Договоры, ваши доли в архиве. А потом смотрите, когда вот вы нажимаете ваше издание, то появляется вот здесь ещё ниже пункты и появляется форма для добавления описания публикации. Переходите к добавлению описания, добавить описание и поехали вперёд описание. Вот это DOI тот самый вот, который вы там получите, и пойдёте дальше там описывать все поля описываете. И эта публикация в конце концов становится вашей публикацией. Но у них там, я вам сказал, они, ребята, не такие, не очень шустрые. На это уходит много времени. Но вот когда я размещал студентам прошлые вот периоды, то я сразу там по 10 штук, по 20 им посылал. Они, значит, где-то неделю вообще ничего не делают. Потом вдруг раз, все 20 это самое, подтвердили. То есть они видно вот заходят, ну, ко мне в кабинет, видят там куча. Всё делают. Потом опять какой-то период ничего не делают. То есть если вот такие большие порции, то они прямо эту порцию всю целиком, так сказать, э-э ну, продвигают на следующий этап. У них там много этапов, меняется статус. Вот. И постепенно потом появляется прямо уже в списке публикаций у вас эта публикация. В РИНЦе, я имею в виду. ResearchGate там всё вообще занимает минуты этот процесс. А так вот и появляются среди публикаций уже здесь вот публикации ваши. Ну у вас должны, должно быть, ну, я не знаю, там, ну две-три публикации должно быть наконец периода вашего обучения. И каждый семестр было бы неплохо там размещать публикацию. И когда будет назначаться стипендия, то это будет учитываться. А в конце, когда будет уже заканчиваться обучение, вы это включите эти публикации, вот эти титульные странички, где публикации из РИНЦа, включите в портфолио или даже целиком саму публикацию туда включите.

5. Использование ПО для Анализа Моделей

Обсуждение функционала ПО

Так, ребят, теперь у нас э-э вопрос такой у меня. Кто из вас посылал мне письма, я вам сказал на занятии будем смотреть? Письма с данными на утверждение темы.

Вы мне говорили, я вам письмо присылал.

Говорил, да? Ну вот, давай теперь мне ссылочку или этот ещё раз пошли, чтобы она наверху там у меня оказалась.

Да-да, сейчас отправлю.

И мы вот будем на занятиях вот этим и заниматься. Тема, что я вам и говорил на занятиях.

Пока не начали, вопрос можно?

Да. Да можно всё время задавать.

Я копался в и нашёл там есть пункт 4.4 решение задач. И написано, что у меня, ну, он в процессе разработки. Это у меня какая-то старая версия или это вообще в целом?

Версия 27.10 должна быть. Вот эта наверху написана, видишь? На окошке написано 27.10 вот здесь.

Ну, у вас, допустим, она работает?

А какой пункт, скажи? 4.4?

Все пункты у меня, ну...

Нет, некоторые пункты не работают. Есть, есть в системе пункты, я вот в меню там их добавил, а не реализовал. Там есть такие случаи, когда напишут, что этот режим находится в разработке. Это может быть. Ну ты скажи конкретно какой. Вот, допустим, там... Вот здесь вот все работают пункты. Вот.

Ну вот 4.4.7 у меня, например.

4.4... Да, да, написано. 4.4.6 написано, да. Это может быть, потому что я же системой занимаюсь. Вот, то есть она находится в процессе разработки.

Я просто думал, у меня какая-то версия...

Нет, нет, нет. Вот, значит, дата вот последней версии, которая сейчас находится в общем доступе, 27.10. Я над ней работаю. У меня и версия есть и более поздняя. Вот, там корректирую, там добавляю что-то, совершенствую. Но, к сожалению, очень большая нагрузка, и я не могу, так сказать, всё это сделать. Когда появятся какие-то окошки мне, тогда я сделаю и размещу. А так вот во время перерывов я еле-еле успеваю там воды попить, грубо говоря, и потом дальше опять вперёд. Поэтому давайте сейчас присылайте данные свои. То есть там может вполне быть, что чего-то и нету.

6. Анализ Студенческих Работ

Пример 1 (Александр Перков): Оружие в игре

Вот София ИТ-2003. А у нас группа вообще-то 2002 сейчас. Так, значит, э-э, Саша Перков. Это...

Это я, да-да.

2002. Вот. В принципе, то, что я вам показываю, ребят, я вам не запрещаю и самим делать. И вот в этой инструкции э-э написал там в инструкции, что вы мне присылаете свои данные, модели. Там даже написано модели. И э-э показываете. Я я смотрю на них, насколько они хорошие эти модели, и э-э мы их утверждаем. Вот смотрите, э-э примерную тему, э-э утверждение возможно только в том случае, если модель получается достаточно достоверная или хотя бы разумная. Ну то есть, что это значит? Это значит, что вы можете взять вот сделать то, что сейчас я делаю, и прислать мне форму, которая подтверждает, что модель разумная, понимаете? Что она нормально получилось. И всё.

Так, значит, значит, теперь давай объясняю. Что это такое вообще? О чём речь тут идёт? Что это за задача? Значит, когда вот так получается, то я советую э-э ячейки эти вот, наименования ячеек делать вертикально. Посерединке. Тогда получается более компактная форма. Вот, записываем. Это у тебя оружие, я так понял, да?

Ну да, да.

А что за оружие? Где там оно, это оружие?

Ну вот первый столбец.

Ну я вижу, что там названия различных этих автоматов там, да?

Ну да, это из игры, да.

Вот. И что мы хотим узнать на основе чего? Данных 45. Значит, в числовых колонках пробелы недопустимы. Там должны быть нули. Если данных нет, может быть, что данных нет вполне. Вот. Ну тогда в числовых колонках мы ставим вместо пробелов ставим нули. Если, вернее, вместо отсутствия данных, ставим ноль. Это будет игнорироваться. То есть при обработке эта колонка будет считаться числовой, но считаться будет, что там нет данных, она не будет никак учитываться. Вот. Ну и скажи мне, значит, э-э, так, это что-то у нас такое? Название оружия, тип оружия. Какие-то характеристики, видимо, оружия, да?

Ну да, там в игре разные характеристики есть, и вот хотелось их сравнить.

Вот. Тогда надо вот что сделать. Тогда мы должны добавить колоночку вот сюда. И в эту колоночку, ну, можно скопировать просто даже, скопировать название оружия. А классификационные шкалы сделать э-э наименование оружия и тип оружия. И всё. Тогда мы сможем сравнить. Значит, здесь их будет не очень много, 45, тут там будет там, скажем, 50-70. Это вполне, так сказать, приемлемо. Смотрим, какие у нас номера классификационные шкалы описательные. Значит, вторая, третья - это классификационные шкалы, четвёртая по двадцать первая - описательная шкала. Вторая, третья - классификационная, четвёртая по двадцать первая, да, так я понял, описательная. Да. Вот. И мы сейчас сможем создать образы этих видов оружия. Конкретных прямо видов оружия и ещё к тому же и категорий оружия, типов. А можно оставить только категории как классификационную шкалу, это вообще убрать. Вот, как хотите. Но, по-моему, так нормально получится. По размерности, если смотреть. 2 3 2 2. Что мы делаем? Какой у нас тип файла экселевского? Новый Excel. Новый. Со второй по третью классификационные шкалы, с четвёртой по двадцать первую описательные. Адаптивные интервалы. С ними всегда получается более достоверно. Ставим птичку, видите, здесь вот переключатель. Нули и пробелы считать отсутствием данных. То есть это вот те самые пробелы. То есть вообще можно считать, можно считать, что вообще там ничего нету у вас. Но если ничего нету в текстовой шкале, то это там пробел. А если ничего нет в числовой шкале, тогда ставится там нолик. Выходим на создание модели. Смотрим, что у нас получилось по наличию данных. Во-первых, матрица получилась 53 на 56. Это хорошая такая маленькая квадратная матрица. Почти что квадратная. Вот. То есть небольшой размерности. Для системы это, так сказать, она это хорошо съест. Теперь смотрите, значит, это вот сами названия конкретных видов оружия. А это вот их там 44. А это название типов, их там девять. И потом характеристики идут. Вот идут характеристики. И каждого вида, в каждый диапазон числовой 12, 13, 15 наблюдений. Это очень неплохо. Это уже как бы случайно не будет, значит, тогда. То есть мы можем сказать, что едва ли там у нас какие-то получатся выводы, едва ли они будут случайные эти выводы, ж понимаете? Когда у нас по 15 наблюдений, вот, то это уже определённая статистика. Дальше, что мы делаем? Выходим на создание модели. Ну это, слава Богу, недолго происходит. Потому что данных вы взяли не 9 ГБ. Уже сообразили, что не надо 9 ГБ брать. И потом мы создаём модели на центральном процессоре и оставляем 30% наиболее значимых результатов. А система сама может определять, насколько какие результаты значимы, насколько они достоверны по уровню сходства. Мы уже много раз видели, что когда уровень сходства увеличивается, возрастает, то и достоверность, э-э, то есть решение тоже возрастает. Практически уровень сходства хорошо отражает степень достоверности, долю истинных решений. Модели уже все созданы. Сейчас идёт их идентификация. Ну говорит, что 17 часов 54 секунды закончен. Передумала, написала 59 секунд закончено. Ну на самом деле, там и будет 59. Потому что в конце там меньше операций, чем в середине. АСК анализ оружия в игре, как называется? Сейчас я в чат напишу. Ничего себе название. Ну здесь в чат писать это, знаете. Здесь-то чат дурацкий же. PUBG. Вот. И смотрим достоверность. Ну, получилось так, скажем так, не великолепно, но более-менее. Я вам скажу так, что всегда получается более-менее, как правило. Редко бывает не совсем проблемы какие-то. Особенно если вы ставите там 30%, оставить наиболее значимых результатов. Ну здесь как-то не особо, между действительно. Так, а здесь? Ну видно, что при повышении уровня сходства доля истинных решений растёт, но как-то она растёт не очень ровно. Тут нет. А здесь даже не очень хорошо вообще. Так. Значит, давай тогда вот что сделаем сейчас. Возьмём и пересмотрим. Не будем конкретное оружие указывать вот это. А только вот будем тип оружия определять по параметрам вот этим. Как ты на это смотришь? Ну хотя тут, конечно, интересно вот это всё. 4 2 2 1. Можно... Ну вот видно, что вот эти вот вверху, они образуют группу, и внизу образуют группу. То есть это как бы противоположные виды оружия, противоположные по их характеристикам. И если мы посмотрим дендрограмму, то там это гораздо нагляднее можно увидеть. Ну придётся подождать немножко. Вот эти данные я послал. Всё. Твоя тема принята, Саш. Александр. Слышишь, нет?

Да, слышу. Спасибо.

Всё, всё, можешь делать описание, получится хорошее у тебя статья, хорошая задача.

Пример 2 (Глеб Клёнс): Квартиры в Москве

Так. Теперь кто ещё?

Я вам прислал работу, ещё раз пришлю.

Да, я вижу. Глеб, да?

Да-да-да. Угу. Квартиры в Москве. А такое точно не было задачи у кого-нибудь, нет?

Нет, это я не посылал ещё до этого.

Потому что я её и помню, да? А том, что что-то не получилось, что-то я не то тебе сказал?

По-моему, по поводу построения этих моделей что-то было.

Потому что, по-моему, вполне такая нормальная задача. Ну что, сейчас давай тогда вот что сделаем. Значит, возьмём, значит, во-первых, смотри. Вот я беру, ставлю клеточку А1, держу клавиши Ctrl и Shift и нажимаю End. У меня выделяется область данных, видишь, ещё как две колоночки. А колоночка К выделяется, вот она. Видишь? Система это вот будет эту область данных анализировать и скажет тебе: "В колоночке К у тебя нет наименования текстового". Ну, в общем, будет мозги морочить тебе. Вот, чтобы этого не было, чтобы было как положено, я создаю пустой лист, вставляю сюда. И на этом всё заканчивается проблема. Делаю классификационную шкалу цена и будем смотреть, как зависит цена от всего остального. Это вообще можно удалить, она не нужна нам. То есть там нарушена вот эта область данных. Ты понял уже, да?

Да-да-да.

Вот. Теперь смотрим, какие у нас тут колоночки по номерам. Вторая - это понятно. С третьей по какую? По десятую. Вторая, вторая, с третьей по десятую. И вводим данные. Данные здесь вроде бы все числовые, я так смотрю. Да, я там переделывал, чтобы всё было с запятыми и прочее. Молодец. Так, это что у нас? Это начало занятия, ребята. Простите меня, сейчас я секундочку отвлекусь. Потому что мне начинаются вот здесь вот занятия. Вы представляете? В этом КГУ. Там расписание не совпадает. Интеллектуальные технологии в науке и образовании. Кошмар вообще. Так, интеллектуальные технологии в науке и образовании. Ладно. Там тем более никого нету. Ну что, значит, вводим мы это вторая, со второй по вторую. В новом виде или да, в старом экселе вводим. Вот. А с третьей по десятую это у нас описательные шкалы. И вот здесь вот смотрите, ребята, интересный момент. Значит, мы классификационные шкалы берём целиком элементы, а описательные шкалы берём элементами, то есть признаками мы считаем слова, которые там есть вот в этих вот полях. То есть слова считаем признаками. Вот Марина. Марин, ты слышишь, что я говорю, нет? Слова считаем признаками. Вот. Это вот здесь вот задано. Слова считаем признаками. И будем определять страну по тем словам, фильмам, там, шоу, которые там крутятся. Какие-то слова будут характерны больше для одной страны, какие-то для другой страны. Вот. H 23.000 строк получится, 555 стран, что ли? Какой-то бред вообще. Ну ладно, сейчас посмотрим. Бред или не бред, сейчас посмотрим. Ну да ладно. Вот. И, значит, выходим на создание модели. Ну там 2 млн, видите, 2 млн... О, ну всё, это нет смысла делать. Это будет у нас, видите, тут нам пишут час, час 40. Но она может быстрее это сделать, а потом сократить прогноз там до 20 минут, до 15. Сейчас посмотрим. Ну да, при таких размерностях это уже... это уже несерьёзно. То есть выходные формы ещё дольше будут считаться. Ну вот, Марина, у тебя получилось, кстати. 25 минут это всё тебе шло, Марина. Марин, ты слышишь меня, нет? Марина не слышит. Я здесь. Алло! Да, это я понял. Я смогла включить его снова. Вот смотри, вот э-э, это твои эти страны и регионы, да? И теперь мы берём и осуществляем синтез модели. Или давай я сейчас тебе это заархивирую, пришлю тебе это. Долго будет. Ну, в общем, понятно, что надо делать. Сейчас мы это и сделаем. Вот. Значит, вот у нас записано со второй по третью. Так, четвёртая, третья и до десятой. Вот это. И мы значит решаем задачу теперь. Запускаем систему в другой папочке, ребята, они спокойно запускаются в разных папочках, не мешают друг другу. Единственное, конечно, работают помедленнее, потому что процессор задействуется. Вот. И, значит, что мы здесь делаем? Значит, берём 2 3 2 2. И указываем какой тип экселя, забыл. Новый. Значит, новый со второй по вторую - это цена. И с третьей по десятую. Равные интервалы берём. Да и разные возьмём. Это получше модели получаются, и они более э-э управляемые, более разумные. Значит, это я, значит, посылаю тебе на почту. Параметры. И посылаю файл тебе на почту. Только не этот. Вот отсюда. Вот. И дальше мы... Ну будет поприличнее. М-м-м-м-м. А почему числовых шкал нету? Название там есть по смыслу. Оно считает, что это не число, представляешь? А что это такое? Это не число. Сейчас посмотрим, в чём дело. И вот это вот смотрим. То же самое. Площадь и числа. А тут видно, в общем-то, тут точки стоят. Или есть числа с точками. А точка - это не числовая данная уже. Доступность, этаж. Вот. Тогда, в общем, более-менее понятно, что надо делать. Сейчас мы это и сделаем. Значит, вот здесь мы поменяем точку на запятую. Восемь раз. И есть ещё один очень интересный вариант. Значит, ну, можно сейчас проверить числа здесь или нет. Каким вот способом? Смотри, Глеб. Я пишу знак равно и сумма. Сумма два раза. Она должна быть до второй строки. А здесь мы видим ерунду полную, что там вообще она не определяет. То есть эта строка, она не считается числовой. Чтобы она стала числовой, есть такой способ. Вот смотри. Беру, значит, переименовываю единичку сюда делаю, добавляю колонку. Беру, выделяю область данных, вставляю в другую как значение. Вот. Вставляю. И потом смотрите, значит, вот там, где у нас числа, мы что делаем? Я пишу равно, восклицательный, то есть единичка, восклицательный знак, это то есть из первой странички я беру эту ячейку B2 + 1 - 1. Так, она вообще её не хочет даже и считать. Обалдеть. Вообще. Тогда надо его в низменный формат преобразовать. Чтобы это сделать, мы создадим файл сейчас вордовский и преобразуем в низменный формат. То есть этот файл, он какой-то вообще дефектный получается. Ну это занимает тут сотни листов, тут 76 листов. Скопировать его теперь. И в Excel перейдём сейчас. Вставим. Давай вот сюда даже. Вообще этот лист нам теперь не нужен. Удалить. Вот сюда и вставим его. Из Ворда. Вот. Теперь ищем лучше всего сразу записывать. Вот. Ну, похоже, теперь на числа хоть, да? Тут какие-то были значки там, не поймёшь там, что оно было. Теперь проверяем, насколько это число. Сумма, сумма. Видишь, с B2 по 2041. Видишь, да? Видно? Да-да-да. То есть всё, это числовая колонка, можно больше тут ни о чём не переживать. Записали его, этот файл. И можно теперь э-э, во-первых, опять тебе его послать. Этот удалить, а этот прикрепить. И теперь можно его и вводить. Вот. И 2 2. Со второй по вторую. В новом виде или да, в старом экселе вводим. Вот. А с третьей по десятую. Вот так вот. И теперь у нас должно быть разумное количество уже этих классов, признаков. Ну вот. Начала она создавать обучающую выборку. И время стало корректироваться, видите? Окончание. Ну я думаю, всё равно неприемлемо, слишком длительно. Не будем же мы полчаса ждать сейчас. Поэтому, Марина, я сейчас тебе пришлю этот файл. А ты с этими параметрами, которые вот мы сейчас задали, они у тебя там в виде этого скриншота там в письме есть. Вот, видишь, вот параметры все есть тут, какие задать тебе. Ты задашь и запустишь. У тебя вот этот процесс пойдёт. Займёт он у тебя, я не знаю, сколько. Ну у нас он где-то вот полчаса может займёт там минут 40. Осталось тут. Ну мы не будем сейчас это ждать. Сейчас будет создана обучающая выборка, а потом нужно будет задать режим... То есть вот это смотри, какие там классы. Что-то 500 классов - это много. У нас же столько стран нету. Нет у нас 500 стран. Да, Николай, привет. Алло, здравствуйте, Евгений Вениаминович. А-а, насчёт вашего письма отправляю. Угу. В плане не получается насчёт ЦНИТа сильно. Угу. Мне как раз таки пути почты спросил. Угу. Сказал вам не писать отправителя. Угу. Нет, там же можете меня написать. А? Меня можно написать, профессора Луценко ЕВ. Отправителя? Да, да, конечно. Профессор Луценко ЕВ. А адрес ваш можно? Ну можно написать Калина 13, кабинет 406. Или кафедра КТС. Написать адрес Калина 13? Ну да, ну да, Краснодар 350, что там, я не помню, 44, по-моему. Вот. Калина 13, кафедра КТС. А, сейчас тут. Так, Краснодар. Вот, и те данные Санкт-Петербург, Московский проспект, дом 192-194, квартира 109, да? Ну там не дом, а корпус. Корпус. Там, по-моему, корпуса. Ну там же это, на Какие вы данные написали, вот я пишу точь-в-точь то, что вы. А, ну то, что да, там вот то, что было, то вот надо и писать. Угу. Ну всё, хорошо. И это, э-э, Коль, этим письмом заказным, потому что документы, и мне прислать на WhatsApp или это фотографию этого чека. Чека? Да. Ну, во-первых, я заплачу, вот, а во-вторых, это э-э там номер для прослеживания, они пусть подчеркнут. Угу, хорошо, я понял. Вот. Всё, сейчас сделаю. Угу. Угу. Давай. Спасибо большое. Там и отзыв, и этот самый, и согласие, да? Я не смотрел. А, ну хорошо, давай. Ну там, я помню вот. Да. Нет, там должно быть две страницы, два два документа. Правильно, да, да, всё правильно, давай. Давай. Спасибо. Вот, ребят, что у нас получается? У нас получается вообще ещё 25 минут она пишет. Ну, конечно, я не думаю, что это стоит ждать. Вот. И меня смущает, почему так много классов получилось. А мы это не узнаем, пока она не закончит. 555 классов. Она ведь, когда мы вот здесь вот задавал я, вот здесь, то я же задавал, видите, описательные шкалы как слова рассматривать, да? А классификационные шкалы целиком элементы. А целиком - это название страны. Что, неужели там 519-520 этих вот... А, хотя там не страна, там, на самом деле, не только страны, там и регионы. Там какие-то есть там регионы ещё там написано. Понятно. Тогда это нормально. Ну всё. Тогда сейчас, э-э, когда-то это, э-э, Марин. Марина! Включай микрофон. Вот смотри, вот у меня считается сейчас ещё 22 минуты. Ага. Я не хочу 22 минуты ждать, понимаешь? Поэтому ты у себя запустишь вот этот процесс и вот досчитается, сформирует э-э шкалы, градации. Можно будет запустить режим 3.5. 3.5 запустишь, он будет считать, он будет сутки считать. Это и больше. Вот. Результаты... Там будет зависимость видна между словами, которые вот в названиях этих фильмов, там, э-э, этих артистов и тому подобное, там всё там аннотации, и регионом в мире, где это. И ты можешь вывести формы по регионам, что вот для этой, для этого региона характерны вот такие вот названия, такие там эти артисты, для этого такие там. И там ещё будут и типы этих шоу, какие, где больше каких смотрят, там, вот всё это будет видно. Но времени займёт много, потому что 6.000 строчек, 22.000 признаков и 500 классов - это всё очень длительные процессы. То есть каждый, каждая строчка будет сравниваться с 550 классами. А чтобы сравнение провести, надо будет 23.000 признаков складывать информацию, чтобы посчитать вообще, о чём речь идёт там. То есть это будет очень длительные расчёты. То есть может там неделю рассчитывать будет, я не знаю, сколько займёт. Ну на графическом процессоре, конечно, нужно считать. Ну если бы вот я считал, то я бы, может быть, не использовал бы не все данные для проверки достоверности модели, а часть данных. Не знаю. Ну давай, пусть она сейчас считает, может, до конца занятия успеет. Сколько там до конца осталось? Ещё довольно долго, ещё полчаса. То есть, возможно, она даже успеет. Тогда я заархивирую эту модель и тебе пришлю. Вот. А сейчас мы можем спокойно в другой папочке открыть систему, запустить и вполне можем там что-то ещё смотреть. Другие модели. Другие процессоры будут нагружены. Поэтому давайте смотреть следующую задачу. Илюха Кина. Ну теперь у нас, я так смотрю, больше таких желающих нет, осталась София. А София здесь есть у нас вообще, нет? Наверное, что её и здесь и нету. Так, это другая группа. Ну я вижу. А в этой группе есть желающие ещё, нет? Что-нибудь проанализировать? Да, есть. Ну давайте. А, Клёнс, я вам прислал работу, сейчас ещё раз пришлю. Да, я вижу. Глеб, да? Да-да-да. Угу. Вот пока она там мучается, мы в это время сейчас у Глеба посмотрим. Ну это я просто просматриваю насчёт того, чтобы там не было пробелов. Ну можно, конечно, вот так было бы сделать, взять просто, если там есть пробелы, заменить их на ну не здесь, а именно вот в области. А тут уже, по-моему, и заменены они на нули и единички. Так. Ну, на будущее я тебе хочу сказать, значит, Глеб, да? Глеб, я хочу тебе сказать, смотри, здесь у нас, видишь, нули и единички, значит, пешком или на транспорте, да? Один пешком, ноль на транспорте. По-хорошему, надо было бы заменить вот это всё. То есть единички заменить на пешком, а нули заменить на транспорте. А, хорошо, сделаю. Вот, допустим, вот так я беру, Ctrl+H. Ну даже давай покажу тебе просто. Вот единичку заменяем на пешком. И всё, понимаешь? Заменить всё. Всё. Потом берём нолик, заменяем на транспорт. Всё. Вот так надо и здесь тоже. Кирпичный, монолитный. Так, H. Единичка. Кирпичный. Кирпич. Монолит. ЖБ. Двойка. Панельный. Тройка. Блочный. Четвёрка. Деревянный. Пятёрка. Соломенный. Как у тех поросят там были. Так, этажность. Этаж, кроме первого и последнего. Иначе. Прикольно. Так. Один. Этаж, кроме первого и последнего. Ноль. Первый или последний. Вот. А потом делаем так: ноль. Первый и последний. О. Классно. И вот так вот. Собственно, и всё. Сейчас эта таблица уже стала гораздо приличнее выглядеть. Вот. А здесь, значит, а здесь мы напишем доступность. Правильно? А здесь напишем проект дома. А это этаж. Всё, ребята. Вот, собственно говоря. Значит, теперь что? У нас у нас записано со второй по третью. Так, четвёртая, третья и до десятой. Вот это. И мы, значит, решаем задачу теперь. Запускаем систему. В другой папочке, ребята, они спокойно запускаются в разных папочках, не мешают друг другу. Единственное, конечно, работают помедленнее, потому что процессор задействуется. Вот. И, значит, что мы здесь делаем? Значит, берём 2 3 2 2. И указываем какой тип экселя, забыл. Новый. Значит, новый со второй по вторую - это цена. И с третьей по десятую. Равные интервалы берём. Да и разные возьмём. Это получше модели получаются, и они более э-э управляемые, более разумные. Значит, это я, значит, посылаю тебе на почту. Параметры. И посылаю файл тебе на почту. Только не этот. Вот отсюда. Вот. И дальше мы... Ну это можно попробовать сделать в системе Идес, а можно попробовать сделать с помощью онлайн-конвертера. Ну давайте сначала попробуем в системе Идес это сделать. Вот сделали. Вот. Ну это занимает тут сотни листов, тут 76 листов. Скопировать его теперь. И в Excel перейдём сейчас. Вставим. Давай вот сюда даже. Вообще этот лист нам теперь не нужен. Удалить. Вот сюда и вставим его. Из Ворда. Вот. Теперь ищем лучше всего сразу записывать. Вот. Ну, похоже, теперь на числа хоть, да? Тут какие-то были значки там, не поймёшь там, что оно было. Теперь проверяем, насколько это число. Сумма, сумма. Видишь, с B2 по 2041. Видишь, да? Видно? Да-да-да. То есть всё, это числовая колонка, можно больше тут ни о чём не переживать. Записали его, этот файл. И можно теперь э-э, во-первых, опять тебе его послать. Этот удалить, а этот прикрепить. И теперь можно его и вводить. Вот. И 2 2. Со второй по вторую. В новом виде или да, в старом экселе вводим. Вот. А с третьей по десятую. Вот это вот. И теперь у нас должно быть разумное количество уже этих классов, признаков. Ну вот. Начала она создавать обучающую выборку. И время стало корректироваться, видите? Окончание. Ну я думаю, всё равно неприемлемо, слишком длительно. Не будем же мы полчаса ждать сейчас. Поэтому, Марина, я сейчас тебе пришлю этот файл. А ты с этими параметрами, которые вот мы сейчас задали, они у тебя там в виде этого скриншота там в письме есть. Вот, видишь, вот параметры все есть тут, какие задать тебе. Ты задашь и запустишь. У тебя вот этот процесс пойдёт. Займёт он у тебя, я не знаю, сколько. Ну у нас он где-то вот полчаса может займёт там минут 40. Осталось тут. Ну мы не будем сейчас это ждать. Сейчас будет создана обучающая выборка, а потом нужно будет задать режим... То есть вот это смотри, какие там классы. Что-то 500 классов - это много. У нас же столько стран нету. Нет у нас 500 стран. Да, Николай, привет. Алло, здравствуйте, Евгений Вениаминович. А-а, насчёт вашего письма отправляю. Угу. В плане не получается насчёт ЦНИТа сильно. Угу. Мне как раз таки пути почты спросил. Угу. Сказал вам не писать отправителя. Угу. Нет, там же можете меня написать. А? Меня можно написать, профессора Луценко ЕВ. Отправителя? Да, да, конечно. Профессор Луценко ЕВ. А адрес ваш можно? Ну можно написать Калина 13, кабинет 406. Или кафедра КТС. Написать адрес Калина 13? Ну да, ну да, Краснодар 350, что там, я не помню, 44, по-моему. Вот. Калина 13, кафедра КТС. А, сейчас тут. Так, Краснодар. Вот, и те данные Санкт-Петербург, Московский проспект, дом 192-194, квартира 109, да? Ну там не дом, а корпус. Корпус. Там, по-моему, корпуса. Ну там же это, на Какие вы данные написали, вот я пишу точь-в-точь то, что вы. А, ну то, что да, там вот то, что было, то вот надо и писать. Угу. Ну всё, хорошо. И это, э-э, Коль, этим письмом заказным, потому что документы, и мне прислать на WhatsApp или это фотографию этого чека. Чека? Да. Ну, во-первых, я заплачу, вот, а во-вторых, это э-э там номер для прослеживания, они пусть подчеркнут. Угу, хорошо, я понял. Вот. Всё, сейчас сделаю. Угу. Угу. Давай. Давай. Спасибо. Вот, ребят, что у нас получается? У нас получается вообще ещё 25 минут она пишет. Ну, конечно, я не думаю, что это стоит ждать. Вот. И меня смущает, почему так много классов получилось. А мы это не узнаем, пока она не закончит. 555 классов. Она ведь, когда мы вот здесь вот задавал я, вот здесь, то я же задавал, видите, описательные шкалы как слова рассматривать, да? А классификационные шкалы целиком элементы. А целиком - это название страны. Что, неужели там 519-520 этих вот... А, хотя там не страна, там, на самом деле, не только страны, там и регионы. Там какие-то есть там регионы ещё там написано. Понятно. Тогда это нормально. Ну всё. Тогда сейчас, э-э, когда-то это, э-э, Марин. Марина! Включай микрофон. Вот смотри, вот у меня считается сейчас ещё 22 минуты. Ага. Я не хочу 22 минуты ждать, понимаешь? Поэтому ты у себя запустишь вот этот процесс и вот досчитается, сформирует э-э шкалы, градации. Можно будет запустить режим 3.5. 3.5 запустишь, он будет считать, он будет сутки считать. Это и больше. Вот. Результаты... Там будет зависимость видна между словами, которые вот в названиях этих фильмов, там, э-э, этих артистов и тому подобное, там всё там аннотации, и регионом в мире, где это. И ты можешь вывести формы по регионам, что вот для этой, для этого региона характерны вот такие вот названия, такие там эти артисты, для этого такие там. И там ещё будут и типы этих шоу, какие, где больше каких смотрят, там, вот всё это будет видно. Но времени займёт много, потому что 6.000 строчек, 22.000 признаков и 500 классов - это всё очень длительные процессы. То есть каждый, каждая строчка будет сравниваться с 550 классами. А чтобы сравнение провести, надо будет 23.000 признаков складывать информацию, чтобы посчитать вообще, о чём речь идёт там. То есть это будет очень длительные расчёты. То есть может там неделю рассчитывать будет, я не знаю, сколько займёт. Ну на графическом процессоре, конечно, нужно считать. Ну если бы вот я считал, то я бы, может быть, не использовал бы не все данные для проверки достоверности модели, а часть данных. Не знаю. Ну давай, пусть она сейчас считает, может, до конца занятия успеет. Сколько там до конца осталось? Ещё довольно долго, ещё полчаса. То есть, возможно, она даже успеет. Тогда я заархивирую эту модель и тебе пришлю. Вот. А сейчас мы можем спокойно в другой папочке открыть систему, запустить и вполне можем там что-то ещё смотреть. Другие модели. Другие процессоры будут нагружены. Поэтому давайте смотреть следующую задачу. Илюха Кина. Ну теперь у нас, я так смотрю, больше таких желающих нет, осталась София. А София здесь есть у нас вообще, нет? Наверное, что её и здесь и нету. Так, это другая группа. Ну я вижу. А в этой группе есть желающие ещё, нет? Что-нибудь проанализировать? Да, есть. Ну давайте. А, Клёнс, я вам прислал работу, сейчас ещё раз пришлю. Да, я вижу. Глеб, да? Да-да-да. Угу. Вот пока она там мучается, мы в это время сейчас у Глеба посмотрим. Ну это я просто просматриваю насчёт того, чтобы там не было пробелов. Ну можно, конечно, вот так было бы сделать, взять просто, если там есть пробелы, заменить их на ну не здесь, а именно вот в области. А тут уже, по-моему, и заменены они на нули и единички. Так. Ну, на будущее я тебе хочу сказать, значит, Глеб, да? Глеб, я хочу тебе сказать, смотри, здесь у нас, видишь, нули и единички, значит, пешком или на транспорте, да? Один пешком, ноль на транспорте. По-хорошему, надо было бы заменить вот это всё. То есть единички заменить на пешком, а нули заменить на транспорте. А, хорошо, сделаю. Вот, допустим, вот так я беру, Ctrl+H. Ну даже давай покажу тебе просто. Вот единичку заменяем на пешком. И всё, понимаешь? Заменить всё. Всё. Потом берём нолик, заменяем на транспорт. Всё. Вот так надо и здесь тоже. Кирпичный, монолитный. Так, H. Единичка. Кирпичный. Кирпич. Монолит. ЖБ. Двойка. Панельный. Тройка. Блочный. Четвёрка. Деревянный. Пятёрка. Соломенный. Как у тех поросят там были. Так, этажность. Этаж, кроме первого и последнего. Иначе. Прикольно. Так. Один. Этаж, кроме первого и последнего. Ноль. Первый или последний. Вот. А потом делаем так: ноль. Первый и последний. О. Классно. И вот так вот. Собственно, и всё. Сейчас эта таблица уже стала гораздо приличнее выглядеть. Вот. А здесь, значит, а здесь мы напишем доступность. Правильно? А здесь напишем проект дома. А это этаж. Всё, ребята. Вот, собственно говоря. Значит, теперь что? У нас у нас записано со второй по третью. Так, четвёртая, третья и до десятой. Вот это. И мы, значит, решаем задачу теперь. Запускаем систему. В другой папочке, ребята, они спокойно запускаются в разных папочках, не мешают друг другу. Единственное, конечно, работают помедленнее, потому что процессор задействуется. Вот. И, значит, что мы здесь делаем? Значит, берём 2 3 2 2. И указываем какой тип экселя, забыл. Новый. Значит, новый со второй по вторую - это цена. И с третьей по десятую. Равные интервалы берём. Да и разные возьмём. Это получше модели получаются, и они более э-э управляемые, более разумные. Значит, это я, значит, посылаю тебе на почту. Параметры. И посылаю файл тебе на почту. Только не этот. Вот отсюда. Вот. И дальше мы... Ну это можно попробовать сделать в системе Идес, а можно попробовать сделать с помощью онлайн-конвертера. Ну давайте сначала попробуем в системе Идес это сделать. Вот сделали. Вот. Ну это занимает тут сотни листов, тут 76 листов. Скопировать его теперь. И в Excel перейдём сейчас. Вставим. Давай вот сюда даже. Вообще этот лист нам теперь не нужен. Удалить. Вот сюда и вставим его. Из Ворда. Вот. Теперь ищем лучше всего сразу записывать. Вот. Ну, похоже, теперь на числа хоть, да? Тут какие-то были значки там, не поймёшь там, что оно было. Теперь проверяем, насколько это число. Сумма, сумма. Видишь, с B2 по 2041. Видишь, да? Видно? Да-да-да. То есть всё, это числовая колонка, можно больше тут ни о чём не переживать. Записали его, этот файл. И можно теперь э-э, во-первых, опять тебе его послать. Этот удалить, а этот прикрепить. И теперь можно его и вводить. Вот. И 2 2. Со второй по вторую. В новом виде или да, в старом экселе вводим. Вот. А с третьей по десятую. Вот это вот. И теперь у нас должно быть разумное количество уже этих классов, признаков. Ну вот. Начала она создавать обучающую выборку. И время стало корректироваться, видите? Окончание. Ну я думаю, всё равно неприемлемо, слишком длительно. Не будем же мы полчаса ждать сейчас. Поэтому, Марина, я сейчас тебе пришлю этот файл. А ты с этими параметрами, которые вот мы сейчас задали, они у тебя там в виде этого скриншота там в письме есть. Вот, видишь, вот параметры все есть тут, какие задать тебе. Ты задашь и запустишь. У тебя вот этот процесс пойдёт. Займёт он у тебя, я не знаю, сколько. Ну у нас он где-то вот полчаса может займёт там минут 40. Осталось тут. Ну мы не будем сейчас это ждать. Сейчас будет создана обучающая выборка, а потом нужно будет задать режим... То есть вот это смотри, какие там классы. Что-то 500 классов - это много. У нас же столько стран нету. Нет у нас 500 стран. Да, Николай, привет. Алло, здравствуйте, Евгений Вениаминович. А-а, насчёт вашего письма отправляю. Угу. В плане не получается насчёт ЦНИТа сильно. Угу. Мне как раз таки пути почты спросил. Угу. Сказал вам не писать отправителя. Угу. Нет, там же можете меня написать. А? Меня можно написать, профессора Луценко ЕВ. Отправителя? Да, да, конечно. Профессор Луценко ЕВ. А адрес ваш можно? Ну можно написать Калина 13, кабинет 406. Или кафедра КТС. Написать адрес Калина 13? Ну да, ну да, Краснодар 350, что там, я не помню, 44, по-моему. Вот. Калина 13, кафедра КТС. А, сейчас тут. Так, Краснодар. Вот, и те данные Санкт-Петербург, Московский проспект, дом 192-194, квартира 109, да? Ну там не дом, а корпус. Корпус. Там, по-моему, корпуса. Ну там же это, на Какие вы данные написали, вот я пишу точь-в-точь то, что вы. А, ну то, что да, там вот то, что было, то вот надо и писать. Угу. Ну всё, хорошо. И это, э-э, Коль, этим письмом заказным, потому что документы, и мне прислать на WhatsApp или это фотографию этого чека. Чека? Да. Ну, во-первых, я заплачу, вот, а во-вторых, это э-э там номер для прослеживания, они пусть подчеркнут. Угу, хорошо, я понял. Вот. Всё, сейчас сделаю. Угу. Угу. Давай. Давай. Спасибо. Вот, ребят, что у нас получается? У нас получается вообще ещё 25 минут она пишет. Ну, конечно, я не думаю, что это стоит ждать. Вот. И меня смущает, почему так много классов получилось. А мы это не узнаем, пока она не закончит. 555 классов. Она ведь, когда мы вот здесь вот задавал я, вот здесь, то я же задавал, видите, описательные шкалы как слова рассматривать, да? А классификационные шкалы целиком элементы. А целиком - это название страны. Что, неужели там 519-520 этих вот... А, хотя там не страна, там, на самом деле, не только страны, там и регионы. Там какие-то есть там регионы ещё там написано. Понятно. Тогда это нормально. Ну всё. Тогда сейчас, э-э, когда-то это, э-э, Марин. Марина! Включай микрофон. Вот смотри, вот у меня считается сейчас ещё 22 минуты. Ага. Я не хочу 22 минуты ждать, понимаешь? Поэтому ты у себя запустишь вот этот процесс и вот досчитается, сформирует э-э шкалы, градации. Можно будет запустить режим 3.5. 3.5 запустишь, он будет считать, он будет сутки считать. Это и больше. Вот. Результаты... Там будет зависимость видна между словами, которые вот в названиях этих фильмов, там, э-э, этих артистов и тому подобное, там всё там аннотации, и регионом в мире, где это. И ты можешь вывести формы по регионам, что вот для этой, для этого региона характерны вот такие вот названия, такие там эти артисты, для этого такие там. И там ещё будут и типы этих шоу, какие, где больше каких смотрят, там, вот всё это будет видно. Но времени займёт много, потому что 6.000 строчек, 22.000 признаков и 500 классов - это всё очень длительные процессы. То есть каждый, каждая строчка будет сравниваться с 550 классами. А чтобы сравнение провести, надо будет 23.000 признаков складывать информацию, чтобы посчитать вообще, о чём речь идёт там. То есть это будет очень длительные расчёты. То есть может там неделю рассчитывать будет, я не знаю, сколько займёт. Ну на графическом процессоре, конечно, нужно считать. Ну если бы вот я считал, то я бы, может быть, не использовал бы не все данные для проверки достоверности модели, а часть данных. Не знаю. Ну давай, пусть она сейчас считает, может, до конца занятия успеет. Сколько там до конца осталось? Ещё довольно долго, ещё полчаса. То есть, возможно, она даже успеет. Тогда я заархивирую эту модель и тебе пришлю. Вот. А сейчас мы можем спокойно в другой папочке открыть систему, запустить и вполне можем там что-то ещё смотреть. Другие модели. Другие процессоры будут нагружены. Поэтому давайте смотреть следующую задачу. Илюха Кина. Ну теперь у нас, я так смотрю, больше таких желающих нет, осталась София. А София здесь есть у нас вообще, нет? Наверное, что её и здесь и нету. Так, это другая группа. Ну я вижу. А в этой группе есть желающие ещё, нет? Что-нибудь проанализировать? Да, есть. Ну давайте. А, Клёнс, я вам прислал работу, сейчас ещё раз пришлю. Да, я вижу. Глеб, да? Да-да-да. Угу. Вот пока она там мучается, мы в это время сейчас у Глеба посмотрим. Ну это я просто просматриваю насчёт того, чтобы там не было пробелов. Ну можно, конечно, вот так было бы сделать, взять просто, если там есть пробелы, заменить их на ну не здесь, а именно вот в области. А тут уже, по-моему, и заменены они на нули и единички. Так. Ну, на будущее я тебе хочу сказать, значит, Глеб, да? Глеб, я хочу тебе сказать, смотри, здесь у нас, видишь, нули и единички, значит, пешком или на транспорте, да? Один пешком, ноль на транспорте. По-хорошему, надо было бы заменить вот это всё. То есть единички заменить на пешком, а нули заменить на транспорте. А, хорошо, сделаю. Вот, допустим, вот так я беру, Ctrl+H. Ну даже давай покажу тебе просто. Вот единичку заменяем на пешком. И всё, понимаешь? Заменить всё. Всё. Потом берём нолик, заменяем на транспорт. Всё. Вот так надо и здесь тоже. Кирпичный, монолитный. Так, H. Единичка. Кирпичный. Кирпич. Монолит. ЖБ. Двойка. Панельный. Тройка. Блочный. Четвёрка. Деревянный. Пятёрка. Соломенный. Как у тех поросят там были. Так, этажность. Этаж, кроме первого и последнего. Иначе. Прикольно. Так. Один. Этаж, кроме первого и последнего. Ноль. Первый или последний. Вот. А потом делаем так: ноль. Первый и последний. О. Классно. И вот так вот. Собственно, и всё. Сейчас эта таблица уже стала гораздо приличнее выглядеть. Вот. А здесь, значит, а здесь мы напишем доступность. Правильно? А здесь напишем проект дома. А это этаж. Всё, ребята. Вот, собственно говоря. Значит, теперь что? У нас у нас записано со второй по третью. Так, четвёртая, третья и до десятой. Вот это. И мы, значит, решаем задачу теперь. Запускаем систему. В другой папочке, ребята, они спокойно запускаются в разных папочках, не мешают друг другу. Единственное, конечно, работают помедленнее, потому что процессор задействуется. Вот. И, значит, что мы здесь делаем? Значит, берём 2 3 2 2. И указываем какой тип экселя, забыл. Новый. Значит, новый со второй по вторую - это цена. И с третьей по десятую. Равные интервалы берём. Да и разные возьмём. Это получше модели получаются, и они более э-э управляемые, более разумные. Значит, это я, значит, посылаю тебе на почту. Параметры. И посылаю файл тебе на почту. Только не этот. Вот отсюда. Вот. И дальше мы... Ну это можно попробовать сделать в системе Идес, а можно попробовать сделать с помощью онлайн-конвертера. Ну давайте сначала попробуем в системе Идес это сделать. Вот сделали. Вот. Ну это занимает тут сотни листов, тут 76 листов. Скопировать его теперь. И в Excel перейдём сейчас. Вставим. Давай вот сюда даже. Вообще этот лист нам теперь не нужен. Удалить. Вот сюда и вставим его. Из Ворда. Вот. Теперь ищем лучше всего сразу записывать. Вот. Ну, похоже, теперь на числа хоть, да? Тут какие-то были значки там, не поймёшь там, что оно было. Теперь проверяем, насколько это число. Сумма, сумма. Видишь, с B2 по 2041. Видишь, да? Видно? Да-да-да. То есть всё, это числовая колонка, можно больше тут ни о чём не переживать. Записали его, этот файл. И можно теперь э-э, во-первых, опять тебе его послать. Этот удалить, а этот прикрепить. И теперь можно его и вводить. Вот. И 2 2. Со второй по вторую. В новом виде или да, в старом экселе вводим. Вот. А с третьей по десятую. Вот это вот. И теперь у нас должно быть разумное количество уже этих классов, признаков. Ну вот. Начала она создавать обучающую выборку. И время стало корректироваться, видите? Окончание. Ну я думаю, всё равно неприемлемо, слишком длительно. Не будем же мы полчаса ждать сейчас. Поэтому, Марина, я сейчас тебе пришлю этот файл. А ты с этими параметрами, которые вот мы сейчас задали, они у тебя там в виде этого скриншота там в письме есть. Вот, видишь, вот параметры все есть тут, какие задать тебе. Ты задашь и запустишь. У тебя вот этот процесс пойдёт. Займёт он у тебя, я не знаю, сколько. Ну у нас он где-то вот полчаса может займёт там минут 40. Осталось тут. Ну мы не будем сейчас это ждать. Сейчас будет создана обучающая выборка, а потом нужно будет задать режим... То есть вот это смотри, какие там классы. Что-то 500 классов - это много. У нас же столько стран нету. Нет у нас 500 стран. Да, Николай, привет. Алло, здравствуйте, Евгений Вениаминович. А-а, насчёт вашего письма отправляю. Угу. В плане не получается насчёт ЦНИТа сильно. Угу. Мне как раз таки пути почты спросил. Угу. Сказал вам не писать отправителя. Угу. Нет, там же можете меня написать. А? Меня можно написать, профессора Луценко ЕВ. Отправителя? Да, да, конечно. Профессор Луценко ЕВ. А адрес ваш можно? Ну можно написать Калина 13, кабинет 406. Или кафедра КТС. Написать адрес Калина 13? Ну да, ну да, Краснодар 350, что там, я не помню, 44, по-моему. Вот. Калина 13, кафедра КТС. А, сейчас тут. Так, Краснодар. Вот, и те данные Санкт-Петербург, Московский проспект, дом 192-194, квартира 109, да? Ну там не дом, а корпус. Корпус. Там, по-моему, корпуса. Ну там же это, на Какие вы данные написали, вот я пишу точь-в-точь то, что вы. А, ну то, что да, там вот то, что было, то вот надо и писать. Угу. Ну всё, хорошо. И это, э-э, Коль, этим письмом заказным, потому что документы, и мне прислать на WhatsApp или это фотографию этого чека. Чека? Да. Ну, во-первых, я заплачу, вот, а во-вторых, это э-э там номер для прослеживания, они пусть подчеркнут. Угу, хорошо, я понял. Вот. Всё, сейчас сделаю. Угу. Угу. Давай. Давай. Спасибо. Вот, ребят, что у нас получается? У нас получается вообще ещё 25 минут она пишет. Ну, конечно, я не думаю, что это стоит ждать. Вот. И меня смущает, почему так много классов получилось. А мы это не узнаем, пока она не закончит. 555 классов. Она ведь, когда мы вот здесь вот задавал я, вот здесь, то я же задавал, видите, описательные шкалы как слова рассматривать, да? А классификационные шкалы целиком элементы. А целиком - это название страны. Что, неужели там 519-520 этих вот... А, хотя там не страна, там, на самом деле, не только страны, там и регионы. Там какие-то есть там регионы ещё там написано. Понятно. Тогда это нормально. Ну всё. Тогда сейчас, э-э, когда-то это, э-э, Марин. Марина! Включай микрофон. Вот смотри, вот у меня считается сейчас ещё 22 минуты. Ага. Я не хочу 22 минуты ждать, понимаешь? Поэтому ты у себя запустишь вот этот процесс и вот досчитается, сформирует э-э шкалы, градации. Можно будет запустить режим 3.5. 3.5 запустишь, он будет считать, он будет сутки считать. Это и больше. Вот. Результаты... Там будет зависимость видна между словами, которые вот в названиях этих фильмов, там, э-э, этих артистов и тому подобное, там всё там аннотации, и регионом в мире, где это. И ты можешь вывести формы по регионам, что вот для этой, для этого региона характерны вот такие вот названия, такие там эти артисты, для этого такие там. И там ещё будут и типы этих шоу, какие, где больше каких смотрят, там, вот всё это будет видно. Но времени займёт много, потому что 6.000 строчек, 22.000 признаков и 500 классов - это всё очень длительные процессы. То есть каждый, каждая строчка будет сравниваться с 550 классами. А чтобы сравнение провести, надо будет 23.000 признаков складывать информацию, чтобы посчитать вообще, о чём речь идёт там. То есть это будет очень длительные расчёты. То есть может там неделю рассчитывать будет, я не знаю, сколько займёт. Ну на графическом процессоре, конечно, нужно считать. Ну если бы вот я считал, то я бы, может быть, не использовал бы не все данные для проверки достоверности модели, а часть данных. Не знаю. Ну давай, пусть она сейчас считает, может, до конца занятия успеет. Сколько там до конца осталось? Ещё довольно долго, ещё полчаса. То есть, возможно, она даже успеет. Тогда я заархивирую эту модель и тебе пришлю. Вот. А сейчас мы можем спокойно в другой папочке открыть систему, запустить и вполне можем там что-то ещё смотреть. Другие модели. Другие процессоры будут нагружены. Поэтому давайте смотреть следующую задачу. Илюха Кина. Ну теперь у нас, я так смотрю, больше таких желающих нет, осталась София. А София здесь есть у нас вообще, нет? Наверное, что её и здесь и нету. Так, это другая группа. Ну я вижу. А в этой группе есть желающие ещё, нет? Что-нибудь проанализировать? Да, есть. Ну давайте. А, Клёнс, я вам прислал работу, сейчас ещё раз пришлю. Да, я вижу. Глеб, да? Да-да-да. Угу. Вот пока она там мучается, мы в это время сейчас у Глеба посмотрим. Ну это я просто просматриваю насчёт того, чтобы там не было пробелов. Ну можно, конечно, вот так было бы сделать, взять просто, если там есть пробелы, заменить их на ну не здесь, а именно вот в области. А тут уже, по-моему, и заменены они на нули и единички. Так. Ну, на будущее я тебе хочу сказать, значит, Глеб, да? Глеб, я хочу тебе сказать, смотри, здесь у нас, видишь, нули и единички, значит, пешком или на транспорте, да? Один пешком, ноль на транспорте. По-хорошему, надо было бы заменить вот это всё. То есть единички заменить на пешком, а нули заменить на транспорте. А, хорошо, сделаю. Вот, допустим, вот так я беру, Ctrl+H. Ну даже давай покажу тебе просто. Вот единичку заменяем на пешком. И всё, понимаешь? Заменить всё. Всё. Потом берём нолик, заменяем на транспорт. Всё. Вот так надо и здесь тоже. Кирпичный, монолитный. Так, H. Единичка. Кирпичный. Кирпич. Монолит. ЖБ. Двойка. Панельный. Тройка. Блочный. Четвёрка. Деревянный. Пятёрка. Соломенный. Как у тех поросят там были. Так, этажность. Этаж, кроме первого и последнего. Иначе. Прикольно. Так. Один. Этаж, кроме первого и последнего. Ноль. Первый или последний. Вот. А потом делаем так: ноль. Первый и последний. О. Классно. И вот так вот. Собственно, и всё. Сейчас эта таблица уже стала гораздо приличнее выглядеть. Вот. А здесь, значит, а здесь мы напишем доступность. Правильно? А здесь напишем проект дома. А это этаж. Всё, ребята. Вот, собственно говоря. Значит, теперь что? У нас у нас записано со второй по третью. Так, четвёртая, третья и до десятой. Вот это. И мы, значит, решаем задачу теперь. Запускаем систему. В другой папочке, ребята, они спокойно запускаются в разных папочках, не мешают друг другу. Единственное, конечно, работают помедленнее, потому что процессор задействуется. Вот. И, значит, что мы здесь делаем?