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

***Российская Федерация***

***Воронин Г.Г. ПИ2104 КубГАУ*** [***glebvr17@gmail.com***](mailto:glebvr17@gmail.com)

[***вашсайт***](mailto:glebvr17@gmail.com)

**210 Лекция №12. По дисциплине Теория информации, данные и знания. 2020-11-17**

**Информация, Знание и Прогнозирование: Анализ данных землетрясений с помощью системы Эйдос**

**Резюме лекции**

Лекция, прочитанная 17 ноября 2020 года профессором Луценко Е.В., посвящена теории информации и данных как знанию. Основное внимание уделяется методологии количественной оценки информации (по Харкевичу) на основе эмпирических данных и ее применению для решения задач идентификации, прогнозирования, поддержки принятия решений и исследования предметных областей с использованием автоматизированной системы когнитивного анализа "Эйдос".

**1. Методология анализа данных:**  
Представлен многоэтапный процесс анализа:

* **Когнитивно-целевая структуризация:** Определение целей исследования, факторов (причин) и результатов (следствий) на понятийном уровне.
* **Формализация предметной области:** Разработка классификационных и описательных шкал и градаций для кодирования исходных данных и формирования обучающей выборки.
* **Синтез и верификация моделей:** Построение статистических (матрицы частот, сопряженности, условных/безусловных распределений) и системно-когнитивных моделей (N1-N7 в системе Эйдос). Оценка их достоверности с использованием F-критерия Ван Рейсбергена и его обобщений.
* **Решение прикладных задач:** Использование наиболее достоверной модели для идентификации, прогнозирования, поддержки принятия решений и исследования закономерностей в предметной области.

**2. Система Эйдос:**  
Подчеркиваются возможности системы Эйдос:

* Работа с различными типами данных (табличные, текстовые, графические).
* Наличие программных интерфейсов для интеграции.
* Автоматизация процесса построения и верификации моделей.
* Преобразование данных в информацию (выявление зависимостей) и информацию в знания (использование моделей для решения задач).

**3. Пример: Анализ данных землетрясений:**  
Демонстрируется применение методологии на примере анализа большой базы данных (119 лет, более 23 тыс. событий) сильных землетрясений по всему миру.

* **Цель:** Выявление факторов, влияющих на характеристики землетрясений.
* **Результаты:**
  + Построены и верифицированы модели.
  + Выявлены закономерности и факторы (широта, долгота, глубина гипоцентра), влияющие на силу и магнитуду землетрясений.
  + Проведена оценка количества информации, содержащейся в факторах о результатах.
  + Сравнение реальной модели со случайной показало наличие значимого "сигнала" (закономерностей) на фоне "шума". Отношение сигнал/шум оценено примерно в 9 дБ.
  + Продемонстрировано повышение достоверности прогнозирования (на 16-17%) по сравнению со случайным угадыванием.
  + Визуализированы зависимости с помощью когнитивных диаграмм и дендрограмм, показывающих взаимосвязи между факторами и результатами, а также между самими факторами.

**4. Практический аспект:**  
Обсуждается вопрос предоставления студентам корпоративных адресов электронной почты для доступа к научным и образовательным ресурсам (например, ResearchGate), подчеркивается важность этого для учебного процесса.

**Вывод:** Лекция демонстрирует практическую применимость теории информации и системы Эйдос для извлечения знаний из эмпирических данных, выявления скрытых закономерностей и решения задач прогнозирования и поддержки принятия решений даже в сложных предметных областях, таких как сейсмология.

**Детальная расшифровка текста**

**Введение**

Здравствуйте, ребята.  
Добрый вечер.  
Здравствуйте.  
Да, здравствуйте, здравствуйте.

**Идентификация лекции**

Сегодня семнадцатое ноября две тысячи двадцатого года.  
Сейчас пятая пара. Что-то я запутался, что ли? Шестая пара.  
Шестая пара. 17:20-18:50.  
По дисциплине: Теория информации, данные, знания.  
Я так понял, что лекция у нас сейчас.  
Это двенадцатая лекция.  
Сейчас я отмечусь везде, ребята, и мы начнем занятие.  
Да, занятие ведет профессор Луценко Евгений Вениаминович.  
Двенадцатая лекция.  
Да, двенадцатая лекция.

**Обзор предыдущей темы: Взаимосвязь материи, энергии, информации**

Мы продолжаем рассматривать вопросы, связанные с взаимосвязью материи, энергии, информации в природе и в обществе.  
Сейчас откроем пособие и перейдем рассматривать дальнейшие вопросы.

**Текущая тема: Методология количественной оценки информации и ее применение**

Так, ребята, ну, пожалуйста, вы попробуйте напомнить мне, на чем мы остановились на прошлом занятии.  
Так, недвоичные меры информации.  
И дальше мы тут рассматриваем пример.  
И пример этот, по-моему, я даже стал вам уже его рассказывать.

**1. Методология анализа данных (по Харкевичу в системе Эйдос)**

Так, ребята, значит, у нас сейчас учебные вопросы:  
Это методика вычисления количества информации Харкевича на основе эмпирических данных и решение на этой основе задач идентификации, прогнозирования, принятия решений и исследования моделируемой предметной области путем исследования модели этой области.  
И здесь у нас идут следующие этапы работы и следующие вопросы.

* Первый вопрос – это когнитивно-целевая структуризация предметной области.
* Второй вопрос – формализация предметной области.
* Третье – синтез и верификация моделей.
* Четвертое – решение задачи идентификации.
* Пятое – решение задач поддержки принятия решений.
* И шестая – решение задач исследования моделируемой предметной области.

Здесь у нас эти вопросы раскрываются в пособии, но я, наверное, буду рассказывать вам все-таки.

* **1.1. Когнитивно-целевая структуризация предметной области**  
  Первый этап – когнитивно-целевая структуризация предметной области. Название такое немножко заумное, но я могу вам сказать, что это название взято из когнитивных, когнитивного моделирования.  
  И смысл его в том, что просто определяется, что мы будем рассматривать при создании модели предметной области, что мы будем рассматривать в качестве причин или факторов, а что в качестве результатов их влияния. То есть это такое этап, на котором мы просто на понятийном уровне определяемся с тем, что мы хотим узнать в результате моделирования, что мы хотим получить. Мы хотим узнать, как вот это влияет на это, например, и все. Вот. Или как вот такие-то свойства влияют на принадлежность или непринадлежность тем или иным категориям.  
  То есть вот на таком уровне ставятся задача. Это понятийная постановка задачи, смысловая. Поэтому она и называется когнитивно-структуризация. То есть когнитивно-целевая. То есть сначала мы определяемся на таком вот понятийном уровне, что мы хотим узнать на основе чего. И потом ставим цель этого. Цель, например, достижение определенных целевых состояний, чтобы объект моделирования, или уже тогда в этом случае его можно называть объектом управления, что нужно сделать, чтобы этот объект управления перешел в заданное состояние. Вот примерно так. Эта задача и решается с помощью модели.
* **1.2. Формализация предметной области**  
  При этом выполняются следующие этапы: этап формализации предметной области, на котором разрабатываются классификационные описательные шкалы и градации, и с помощью них кодируются исходные данные, получается обучающая выборка.  
  Описательные шкалы – это способ описания, способ формального описания либо свойств, их значений, либо факторов, их значений. А классификационная шкала – это способ формального описания состояния объекта моделирования, которым она, то есть, скажем так, категория обобщающая, к которым относятся состояния объекта моделирования.  
  Эти состояния могут относиться по времени к тому же периоду, когда у него наблюдаются определенные значения свойств. То есть одновременно с этим считается. Тогда мы говорим о статической задаче, что объект есть определенные свойства, и он принадлежит к определенным категориям. И динамический вариант, когда на объект действуют определенные факторы, и он переходит в некоторые будущие состояния.
* **1.3. Синтез и верификация моделей (Система Эйдос)**  
  Когда у нас проведена формализация предметной области, то мы имеем все необходимое для начала создания моделей, расчетом статистических моделей и основанных на них моделей знаний.  
  Система Эйдос, которую вот я разработал, у нее есть одно очень важное достоинство, что в ней есть развитая система программных интерфейсов, связывающих ее с различными объектами, внешними по отношению к системе, информационными, различных типов данных: табличные данные, числовые, текстовые, текстовые файловые данные в файлах текстовых, и графические объекты могут исследоваться тоже в разных стандартах JPEG и BMP, представленные.  
  И благодаря тому, что в системе есть эти программные интерфейсы, она очень становится функциональной, удобной. То есть в ней можно очень быстро создавать приложения, без особых таких усилий, вводя в нее исходные данные. И можно на это не очень много времени тратить и сил и рассматривать различные варианты моделей. То есть можно несколько моделей рассмотреть за короткое время, выбрать из них наилучшую и решать в ней задачи.  
  Этап синтеза и верификации моделей. На этом этапе, как уже понятно из самого названия этого этапа, осуществляется разработка моделей, то есть их расчет статистических и системно-когнитивных или моделей знаний.  
  Статистические модели делятся на три группы в системе Эйдос. Это матрица абсолютных частот, которую еще называют матрица сопряженности или корреляционная матрица. Такая матрица, ребята, рассчитывается, ну, в общем-то, практически во всех системах и статистических, и интеллектуальных, практически во всех, за исключением нейронных сетей, может быть. На основе неё рассчитывается матрица условных и безусловных процентных распределений. Условных – условных это означает распределение числа встреч какого-то признака у объектов, относящихся к определенной категории, к определенному классу. А безусловных – это по всей выборке исследуемой.  
  Потом на основе этих моделей, условных и безусловных процентных распределений, рассчитываются вот матрицы системно-когнитивных моделей, которые я назвал N1, N2, N3, N4, N5, N6 и 7, N7. Можно было бы их и больше сделать, но у меня, так сказать, такого желания не возникло по одной очень простой причине: это бы переусложнило бы систему. То есть и так там довольно много всего, и, скажем так, неоправданное усложнение системы было бы. Возможно было бы, и интегральные критерии можно еще добавить. Я вот сейчас знаю еще варианты, какие можно было бы реализовать. Я их не стал реализовать, потому что у нас в форме выходных, где результаты идентификации, так у нас там два окошка справа, соответствующие разным интегральным критериям. А так было бы три окошка. Ну, можно, конечно, и три сделать. Или можно сделать переключатель интегрального этого критерия и смотреть то на одну, то на другую. Может быть, так удобнее было бы.  
  Вот, ну я этого не сделал. Может быть, кто-нибудь из вас в будущих каких-то своих разработках это сделает когда-нибудь.  
  Потом выбирается наиболее достоверная модель в соответствии критериям, F-критериям Ван Рейсбергена и его обобщениями, которые я предложил: нечеткими, мультиклассовыми, инвариантными относительно объемов данных обучающей выборки.
* **1.4. Решение прикладных задач**  
  И потом в этой наиболее достоверной модели решаются задачи распознавания, идентификации, прогнозирования, принятия решений и исследования моделируемой предметной области.  
  Вот, собственно говоря, что делается.  
  И при этом мы находим причинно-следственные зависимости в исходных данных, преобразуем этим самым эти исходные данные в информацию, отражающую уже причинно-следственные зависимости. И потом, зная эти причинно-следственные зависимости между событиями прошлыми и будущими, между факторами и результатами их влияния, мы решаем задачи идентификации, принятия решений, то есть управления, достижения целей и исследования. То есть уже мы информацию преобразуем в знания, когда мы это делаем.  
  Операция преобразования данных в информацию называется анализ данных, а операция преобразования информации в знания называется управление или постановка цели, достижение целей.

**Пример: Анализ данных землетрясений**

Как мы это делаем на конкретном примере, я вам покажу.  
Этот пример очень интересный. Значит, здесь я привел в развитой форме этот пример, чтобы показать вот просто именно действительно развитой форме показать возможности системы Эйдос.

* **2.1. Описание данных**  
  Значит, ребята, смотрите. Значит, мы видим, что у нас есть исходная база данных, включающая сильные землетрясения по всему миру за период 119 лет, с 1900 года по 2019. Почему за этот период? Потому что это исследование, которое сейчас вот я вам показываю, я провел его в начале года. Вот. Еще в начале, да, 2000, нужно сейчас что-то запутался, какого года. Ладно, ну, в общем, в начале года я его провел. А какого, я не помню.  
  Вот, ну раз за 119 лет, то, наверное, все-таки в начале девятнадцатого. То есть в начале двадцатого года. Хотя, хотя. Вот.  
  И этот, эта база данных включает большое число примеров. Ну, здесь вот написано, что 23.655 примеров землетрясений со своими датами, временем, силой. Вот.  
  И, значит, здесь мы что видим в этой таблице? Мы видим 23.655.000 строк, которые здесь их не видно, потому что, естественно, я привел фрагмент. И также видим что? Что первая колонка содержит информацию об источнике наблюдений, о дате фактически. Вторая колонка никак не используется. Это время, да, как бы? Сейчас посмотрим. Да, похоже, что это время.  
  Потом мы у нас есть вот эти желтые колоночки – это характеристики землетрясений: магнитуда, там, сила, в общем, отражающие различные показатели. Вот, результаты воздействия факторов.
* **2.2. Формализация в примере**  
  Затем идут колонки, в нашем случае их 10, которые описывают сами факторы. Колонки с результатами называются классификационными шкалами и градациями, классами. А колонки с факторами называются описательными шкалами с их градациями, значениями факторов или признаками.  
  Данных, тип данных в этих колонках может быть как числовой, так и текстовый.  
  Сейчас попробуем, знаете, что сделать? Попробуем эту таблицу скачать. Последнее занятие, поэтому я могу все это позакрывать. Попробуем эту таблицу скачать по ссылочке. Что-то ее там нету, что ли? Вот это да. Так, 157-е приложение. Так, когда показываешь, ты и узнаешь. 157-е приложение. Ну да. Эта таблица больше мегабайта. Поэтому она, ну это где-то допустимо. Вообще другое приложение. Замечательно вообще. 156, 156. То есть вот здесь у нас опечаточка. Здесь у нас 156. Вот, ну это уже оно. 23.655 строк. Все правильно, да? Видите, ребята? Все, все верно. Вот, то есть это та таблица.  
  Вот, то есть здесь у нас была неточность.  
  И вот у нас, соответственно, на этапе формализации предметной области созданы описательные шкалы и классификационные шкалы. Вот таблички, которые приводят эти шкалы. Просто перечень этих шкал.
* **2.3. Синтез и верификация моделей в примере**  
  На этапе, да, ребята, что-то хотите сказать, говорите. Может, какие-то вопросы у вас возникли там или еще что-то? Или просто микрофон включили нечаянно?  
  Этап формализации предметной области. На этом этапе разрабатываются классификационные описательные шкалы и градации и обучающая выборка. Создаются вот эти параметры, которые здесь мы видим: классификационные шкалы с третьей по седьмую, описательные с восьмой по семнадцатую. Тип данных новый Excel, адаптивные интервалы. И задаем по два интервальных значения классификационных шкал и по 12 в описательных. У нас получается огромное число наблюдений на диапазон, около 5.600 наблюдений на диапазон. 49 секунд идет формирование классификационных и описательных шкал и градаций и обучающей выборки. Формируются следующие классификационные, следующие описательные шкалы и градации.  
  Мы видим, что очень большое число наблюдений на числовой диапазон попадает, 5.600. А на описательных шкалах 947, около тысячи. Это говорит о том, что эти результаты, которые здесь будут получены в этой модели, они весьма убедительны. То есть они, как скажем, удовлетворяют самым придирчивым требованиям статистики. Это обучающая выборка. Значит, здесь наши 23.000 с хвостиком наблюдений уже в закодированном виде.  
  И дальше мы рассматриваем, как осуществляется синтез и верификация моделей. Ну я про это рассказывал вам. Мы смотрим, какое значение фактора, сколько раз встретилось у объектов той или иной категории. Вот, объектами различных категорий у нас в данном случае являются землетрясения различной силы, залегающие на разных глубинах гипоцентра этих землетрясений.  
  И мы на основе вот этой матрицы частот рассчитываем матрицу условных безусловных процентных распределений. И затем матрицу, которая отражает, какое количество информации содержится в тех или иных факторах о том, что наступят то или иное землетрясение с теми или иными характеристиками.  
  Это матрица хи-квадрат. Здесь вообще мы видим большие величины, которые довольно сильно друг от друга отличаются. Положительные черного цвета, красные отрицательного цвета.
* **2.4. Оценка достоверности и сравнение моделей**  
  Смотрим на то, какая у нас достоверность моделей. Значит, я могу вам сказать, что эта достоверность, которая у нас получилась, в модели, согласно критерию L1, модели NF3 у нас получилось 0,642. Это, вообще-то, неплохо для такой предметной области, как прогнозирование землетрясений. Это вообще даже неожиданно, что получился такой результат.  
  Вот такое частотное распределение получилось. Вы уже знаете, что оно означает, что у нас истинных значений положительных решений всегда больше, чем ложных положительных решений. И отрицательные решения то же самое, тоже всегда их гораздо больше истинных решений, чем ложных.  
  И мы видим, что уровень сходства, при росте уровня сходства закономерно растет доля истинных решений. Это означает, что уровень сходства является адекватной мерой истинности решения и степени уверенности в решении системы.  
  Очень странно вообще сам факт того, что обнаружены эти закономерности. Это, вообще-то, большое достижение, потому что не редко кому удавалось эти закономерности обнаружить. Ну вот наша технология это позволяет сделать.  
  Ну, что это дает? Значит, эта технология позволяет повысить достоверность прогнозирования по сравнению со случайным угадыванием для положительных решений на 17%, для отрицательных на 16%. В среднем на 16,7%.
* **2.5. Анализ результатов и выводы по примеру**  
  Вот, собственно, мы выявили закономерности между факторами и результатами их действия.  
  Теперь возникает вопрос: а какая доля шума в наших исходных данных? То есть там есть и достоверная информация о моделируемой предметной области, и шум. Хотелось бы сравнить и понять, значит, сколько там отношение сигнал/шум, скажем так. Какое там отношение сигнал/шум?  
  Для этого мы создаем точно такую же модель по всем характеристикам, столько же объектов обучающей выборки, столько же классов, к которым относятся один объект, столько же у него признаков, сколько в среднем у этих объектов обучающей выборки. В общем, по всем параметрам размерности мы создаем точно такую же модель, совершенно аналогичную.  
  И решаем при тех же параметрах модели, решаем те же самые задачи, то есть создаем модели, проверяем их на достоверность. Здесь описано, как мы создаем эту случайную модель. Для этого предназначена лабораторная работа 201, которая создает сама, измеряет характеристики текущей модели и создает случайную модель, полностью аналогичную текущей по всем ее параметрам.  
  И дальше мы создаем при тех же параметрах уже случайную модель. Это занимает чуть-чуть больше времени. Там было 58 минут 57 секунд, а здесь час 57 секунд, да, на всяком минут дольше, вот где минут.  
  Мы использовали генератор фон Неймана псевдослучайных чисел, который я немножко улучшил путем использования внешнего источника энтропии, в качестве которого я использовал компьютерный таймер. И исключил циклы в этом генераторе фон Неймана, в нем были циклы. А теперь получилось так, что все числа всегда отличаются, когда бы мы не проводили исследования.  
  Сравнение частотных распределений истинных и ложных положительных и отрицательных решений в реальной и случайной модели будет сейчас проведено. Сравнение достоверности положительных и отрицательных решений в реальных случайных моделях и сравнение информационной мощности реальной случайной модели и оценка отношения полезного сигнала к шуму.  
  Значит, смотрим, ребята. Значит, у нас в реальной модели различие между истинными и ложными решениями более сильное, чем в случайной модели. Что вообще странным является, что меня удивило, что она в случайной модели тоже нашла закономерности. Но они не так ярко выражены, как в реальной модели. Из этого можно сделать вывод, что реальная модель содержит не только шум, но и истинный сигнал. А случайная модель оценивает именно уровень шума. То есть если мы вычтем уровень шум из реальной модели, вот тот, который здесь мы видим, то мы как раз получим и реальный сигнал.  
  Значит, вот сравнение проведено со случайным угадыванием для положительных и отрицательных решений в реальной модели и случайной. И мы видим, насколько повышается достоверность прогнозирования по сравнению со случайным угадыванием. Мы, что самое странное, что в случайной модели мы тоже повышаем достоверность прогнозирования. То есть модель системы Эйдос очень мощная, она хорошо работает даже с шумом, даже в шуме находит закономерности.  
  Тем не менее, мы видим, что довольно существенно отличается достоверность прогнозирования по сравнению со случайным угадыванием в реальной модели и случайной. И разность составляет около восьми с хвостиком процентов. Это означает, что у нас в реальной модели есть сигнал полезный, и он где-то примерно в два раза, ну, в общем, они примерно равны, шум и случайный сигнал. Сопоставимы по величине.  
  То есть это означает, что наша гипотеза подтверждена о том, что факторы, которые мы используем в модели, реально оказывают влияние на результат.  
  Значит, здесь формулы, которые я не буду сейчас вам объяснять, но это формулы для оценки отношения сигнал/шум, которые, в общем, общеизвестны, общеприняты.  
  И мы можем оценить мощность реальной модели по сравнению случайной и отношение сигнал/шум.  
  У нас получается, что влияние факторов, которые мы исследуем, более чем в восемь раз выше влияния шума, что с очень высокой убедительностью подтверждает предложенную нами гипотезу о том, что реально существует влияние этих факторов на результат.  
  Значит, у нас получается отношение сигнал/шум 9 децибел. 9 децибел – это достаточно тихое сообщение. Однако его вполне можно услышать и понять. С помощью сообщения примерно такой громкости врачи проверяют слух. Вообще, слышите вы? Он говорит: "Здравствуйте". Вы говорите: "Да, здравствуйте" там. Ну, короче, проверяют, насколько вы слышите. Информацию об этом можно получить, например, по этой ссылке.
* **2.6. Использование модели для решения задач**  
  Значит, сравнение реальной и случайной модели позволяет сделать убедительные выводы о том, что мы действительно обнаружили влияние этих факторов на сейсмическую активность Земли. И не просто ее обнаружили, а еще описали, как они воздействуют, какие они вызывают результаты.  
  На основе этого можно решать задачи прогнозирования. Для этого используются два интегральных критерия. Один из которых представляет собой скалярное произведение вектора классов и вектора состояния объекта, что в координатной форме выглядит следующим образом. Этот интегральный критерий является неметрическим интегральным критерием и одновременно является определением шума. Это значит, что при увеличении объема выборки, если эта выборка случайная совершенно, только из шума состоит, то эти интегральные критерии, и как и сами частные критерии в матрицах модели, так и интегральные критерии стремятся к нулю.  
  Вот. Если же в исходных данных есть полезный сигнал, то такой интегральный критерий приводит к тому, что отношение сигнал/шум растет при увеличении объема выборки. При неограниченном увеличении объема выборки шум стремится к нулю, и получается, что этот интегральный критерий тоже подавляет шум и получается все лучшие результаты. Но мы видим, что уже даже при 23.000 примеров совершенно очевидно влияние тех факторов, которые мы исследовали на моделируемую предметную область.  
  То есть если кто-то говорит, что "а, это не влияет", то это голословно. Они, эти факторы влияют, и мы это обнаружили и измерили.  
  А это второй интегральный критерий, который является тоже скалярным произведением, но только уже стандартизированных векторов.  
  Дальше мы присваиваем наиболее достоверной модели статус текущий и решаем в ней задачу идентификации, прогнозирования. Это заняло у нас 2 секунды на 23.000 примеров распознаваемой выборки. Файл распознаваемой выборки имеет точно такую же структуру, как файл исходных данных, но может не быть классов указано. И при его вводе в систему мы должны вот здесь поставить птичку не формализация предметной области, а генерация распознаваемой выборки.  
  В результате мы осуществляем ввод распознаваемой выборки за 28 секунд и распознавание прогнозирования проводится в режиме 4.1.2 на центральном процессоре.  
  Значит, тот период, когда писалась эта статья, модуль расчетов на графическом процессоре находился в отладке. Сейчас он, эта отладка завершена, он работает, но выдает одну постоянно одну и ту же ошибку при расчете модели ПРЦ-2. Я в системе Эйдос на центральном процессоре эту ошибку исправил. Почему не на графическом? Потому что модуль распознавания и синтеза моделей распознавания на графическом процессоре разрабатывал Дима Бандыкин с Беларуси. У них сейчас там, сами знаете что, и он сейчас, ну, практически не может заниматься работой.  
  Вот мы примерно 5 минут провели распознавание в этой с помощью центрального процессора, подчеркиваю. То есть это вполне такие разумные сроки, вполне приемлемые. И видно, что процесс прогнозирования занял всего 5 минут. Вот, на графическом процессоре этот процесс занял бы несколько секунд.  
  Ниже приведены экранные формы с результатами решения задачи прогнозирования.  
  Значит, результаты эти совпали с опытом. При этом мы видим, какие модели и какие результаты достоверности, оценка достоверности, какие классы лучше распознаются, какие хуже.  
  Значит, вот мы видим, что вот эта вот глубина землетрясения глубокая, она обусловлена 13% суммарной значимости. А вот широта максимальная всего 6%. То есть получается, что в два раза ниже. То есть степень обусловленности различных характеристик землетрясения тоже различна. Некоторые более жёстко обусловлены, некоторые менее жёстко. Разница в два раза. А по влиянию факторов в 41 раз разница. То есть есть факторы, которые в 41 раз сильнее, чем другие влияют. Ну это, ребята, серьёзно. Серьёзная разница.

**Визуализация результатов**

Значит, здесь я указываю на то, что система Эйдос довольно широко распространена во всём мире. По ней издано много публикаций, монографий, получено много патентов. Тем не менее, некоторые вещи, которые делает система, малоизвестны. Что меня лично меня удивляет и досаду определённую. Пусть они пишут, что это возможно будет в будущем. А это возможно было вот это то, что здесь нарисовано, это было возможно 30 лет назад. 30, ребята, лет назад это всё рисовалось уже в досовской версии системы Эйдос. О чём можно посмотреть в моих работах того времени.  
И такие формы тоже были. Я в системе, которая сейчас у нас под Windows реализована, реализовал примерно треть того, что было в досовской версии системы Эйдос. Треть форм примерно. Там их было гораздо больше этих форм. Но я реализовал только те формы, ребята, слушайте внимательно, и режимы только те и формы, которые хоть раз использовались. В досовской версии системы Эйдос там было очень много понаделано всего, что никогда никем не использовалось, никогда в жизни, ни разу. Только я это разработал, провёл отладку, и больше это никому было не нужно никогда. Я когда делал под Windows версию, то я эти режимы не стал реализовать, которые никогда никем не использовались.  
Значит, это вот у нас, ребята, когнитивная функция, которая показывает, как влияют значения фактора, который здесь описан, на результаты, характеристики землетрясений. Видно, что совершенно однозначно влияет. Прямо определённо. Причём некоторые виды влияния очень, скажем так, почти жёсткие, детерминистское влияние. Вот где красный цвет и синий яркий – это жёсткое влияние, это жёстко обуславливает это, это жёстко препятствующее. А здесь мягкое влияние, но не неоднозначное, но тоже есть.  
Вот мы видим, как влияют различные значения факторов. Видим степень детерминированности и ценность значения факторов для решения задач прогнозирования и принятия решений моделирования. Мы видим, что наиболее значимый фактор имеет 5,7 суммарной значимости, значимость. 25 факторов обеспечивают 50% суммарной значимости, а 50 факторов обеспечивают 70% суммарной значимости. А последний фактор, который наименее значимый, он имеет значимость 0,139. 0,139. Вот разделим мы сейчас 5 737 разделим на 0,139. 41 раз, ребята. То есть наиболее значимый фактор имеет силу влияния в 41 раз больше, чем самый слабо влияющий фактор.  
И мы можем посчитать среднюю значимость шкал. И мы видим, что наиболее значимой шкалой является пятая, она обеспечивает четверть суммарной значимости. А самая мало значимой является, значит, шкала номер девять, и она обеспечивает в пять раз меньшую значимость силу влияния.  
И разные факторы тоже, то есть разные состояния объекта моделирования тоже они обусловлены с разной степенью жёсткости. Есть которые жёстко обусловлены, есть которые более мягко обусловлены.  
Мы это можем визуализировать также и в виде дендрограмм. Ну здесь такая длинная дендрограмма, 120 у нас тут значений факторов. Это уже показывает наглядной форме, что все значения факторов делятся на две группы, которые здесь обозначены красным и синим цветом. И внутри них тоже просматриваются кластеры. Вот эти группы. То есть мы видим, какие значения факторов имеют сходное влияние. Мы это можем визуализировать также и в виде свод-диаграмм. Как влияет тот или иной фактор на результаты. И также в виде нейронной сети. Вот верхний рисунок у нас 98% нейронной сети показано. А нижний – 38%. И вполне, так сказать, чётко видно, что какие-то значения факторов прямо определённо влияют на на некоторые характеристики землетрясений. Причём влияют как положительно, так и отрицательно. То есть некоторые способствуют, некоторые препятствуют.  
График межклассовых расстояний. И смотрим, как влияют, как связаны сами факторы друг с другом. Есть ли факторы, которые влияют сходно или нет? Ну матрица сходства значений факторов, она значительно большую размерность имеет. И поэтому здесь только фрагменты приводятся. Чем матрица сходства классов. 120 строк, 120 колонок уже здесь на не поместишь уже её сюда в книжку. Ну её можно считать, открыть в Экселе.

**Практический аспект: Корпоративные адреса электронной почты**

Ну, в общем, когда землетрясение глубокое, то его интенсивность низкая. Ну это интенсивность имеется в виду на поверхности Земли.  
Кто-нибудь из вас уже зарегистрировался в ResearchGate, ребята, скажите? Кто-нибудь из вас в ResearchGate зарегистрировался? Ну кто да, тот скажите, что да, зарегистрировался. О! Степан зарегистрировался и Марк. Ну, молодцы. Так. Сейчас я вас найду. Как ты написал, Марк, фамилию латинскими буквами, как ты написал ее? О, господи. Ну надо свою писать фамилию. Это же социаль... ну ты как в этом, как будто это одноклассники, да, или кто там у тебя? Что-то я такого нету. Находит.  
Вот. Смотрите, кто у нас тут есть, а? Это же Степан. Вот это вот, знаете, что значит? Вот я сейчас здесь клацаю, смотри. Смотрите. И я становлюсь человеком, который следит за твоей активностью, Степан. Вот. После этого я могу тебе писать сообщения.  
Вот. Теперь, когда ты сделаешь описание, покажешь его мне, чтобы я его утвердил, чтобы там не было, ну, скажем так, лажи какой-нибудь, чтобы было, ну, качество определённый уровень качества. Я тебе скажу: "Всё, теперь размещай". И можешь мне, кстати, вот здесь писать в сообщениях, потому что я тут бываю. Вот, и добавлять будешь вот так: раз, препринт. Всё на английском языке надо писать. Можно пользоваться переводчиком. Указываешь, значит, я делаю так: сначала размещаю вордовский файл, получаю на него DOI, потом вставляю это DOI в вордовский файл, делаю PDF и потом размещаю PDF-файл. Для чего я так делаю? Дело в том, что когда вы выходим на какую-то публикацию, смотрим, если там PDF-файл задан текущим, последний файл PDF был, то сразу же моментально этот файл открывается для чтения. А если вордовский файл, то он э-э открывается только первая страничка или вообще он не открывается, но можно его скачать. Вот. А если есть вот первый файл вордовский, а потом PDF, то PDF сразу открывается для чтения, а а вордовский можно скачать несложно.  
Ты интересуешься алгеброй, компьютерными коммуникациями, алгоритмами, базами данных. Ну вообще круто. Вот. Ну всё, молодец. Молодец. Вот, а Марк, тебе что-то такое у тебя не получилось, мне кажется. Тебе надо попробовать ещё раз зарегистрироваться. Я даже не очень уверен. Я забыл, какую почту указал туда. А! Господи. Ну надо корпоративную указывать. Вот. э-э Ну, в общем, короче говоря, тебе могут не дать второй раз. Вот в чём дело. Какие-то у них могут быть свои там заморочки. Ну, короче говоря, тебе надо зарегистрироваться, только надо по-нормальному написать там свои имя, фамилию по-английски.  
О! Значит, смотри, смотри, Александр. Вот я показываю тебе. Вот эта папочка, где разработки ведутся. Там разные языки используются. Ты ж видишь, да, там, наверное, на экране. Вот я сейчас беру, смотри, вот исходный текст. Вот этот исходный текст вчерашний. Даже сегодняшний, сегодняшне утренний. Вот. Весь этот текст вот я беру и компилирую. Возникли проблемы с компиляцией, но я их обошёл простым способом, хотя немножко жалко было. Ну я заархивировал предыдущую версию. В смысле, зафиксировал. Вот, и потом продолжил с такой версией, которая спокойно компилируется. Но это было некоторой ценой, которая мне лично вот обидно, что это пришлось сделать. Значит, что пришлось сделать? Я взял одну функцию, которая не связана с другими, выкинул из системы. Очень хорошая функция, интересная, ценная. Я даже скажу какая. Это тот режим, который делает вот те расчёты, которые я сейчас вам показывал. Но только я сейчас то, что я показывал, это стандартная функция системы Эйдос используется. А там был сделан режим прямо специально для таких вот расчётов, связанных с землетрясениями. Очень хороший режим. Сейчас я вам его покажу. Книжку про это покажу. Ну да, я даже вот здесь вот я сейчас так вот найду. Сейчас, э-э Вот она. Вот, смотрите. Вот он этот модуль. Экзем-модуль 20 МБ. 17.11.2020 19:48. Видишь, да? Вот он этот модуль. Ну то есть как бы выкрутиться-то несложно. Но вот такой ценой, что модуль, который рассчитывал, делал расчёты вот для этой монографии, он оказался в одной из версий, которая зафиксирована, и, в общем, она эта версия не меняется уже. Вот. И вот мы можем это почитать. Идём, смотрим. А можем просто скачать. Можем скачать просто и потом читать. Вот. И вот эта книжка, ребята. Это монография, изданная в этом году. Очень такая серьёзная. Не то, что я сейчас вам сегодня рассказывал, это глава из этой монографии, моя. Но там есть и другие главы, которые не я писал, но я делал программы для их для того, чтобы это можно было сделать эти расчёты, которые здесь описаны. Ну эти программы – это система Эйдос и ещё некоторые режимы дополнительные, которые просто были в меню. Я их изменил, убрал.  
Значит, э-э, ну, старосты присутствуют, ребята? Кто тут у нас староста? Присутствует, да? Да, да, присутствует. Ребята, значит, слушайте внимательно. Делайте список группы, распечатываете и относите в центр информационных технологий в кабинет Алексея Семёновича Креймеру. И пишете, э-э, в произвольной форме перед списком группы пишете там служебная записка директору центра информационных технологий Алексею Семёновичу Креймеру, он доцент, доценту Алексею Семёновичу Креймеру, служебная записка. Прошу присвоить или выдать адреса корпоративной электронной почты студентам группы такой-то. И список студентов. И поле, ребята, оставьте, список студентов, просто 1 2 3, фамилия, имя, отчество. А сбоку оставьте достаточно поле, чтобы можно было вручную вписать адрес электронной почты. И потом внизу подписывайтесь: староста группы и своя подпись и фамилия, инициалы. И относите Алексею Семёновичу Креймеру. Он, я так понял, что это сейчас уже является основанием. То есть он как бы пошёл навстречу, я так сказал. Ну, наверное, в общем, короче, он понял, что это нужно действительно. Потому что постоянно идут обращения, не только вот в связи с тем, что я там говорю вам, давайте там в ResearchGate регистрироваться, не только я такой, а вообще это нужно, многие системы требуют этого адреса. Вы же знаете, да, вот и Марк это вот сейчас написал. Вот. Короче говоря, делайте вот такой списочек, относите. И потом он вам, вам скажут прийти там через 2 дня, к примеру. И возьмёте этот списочек и сообщите, отфотографируйте, пошлёте в группу свою Ватсаповскую по своей группе, и всё, понимаете? Значит, там адреса будут очень просто устроены. То есть там фамилия идёт в транслитерации латинскими буковками, потом точка, первая буква имени, лягушка, потом edu.kubsu.ru. У вас у всех вот такие появятся адреса. Такие же, как в КГУ, только там не Кубсау, а Кубсу. Это джеймеловские адреса электронной почты. Домен, то есть они позволяют Gmail позволяет домен делать своего вуза. И для ResearchGate этого должно быть достаточно для регистрации. Если будут какие-то проблемы, опять же к ним в ЦИТ можно обратиться. Да, Ремёр, что вы хотели сказать? Это важная информация. Сообщите её всем, кого там знаете, грубо говоря, и кого не знаете.

**Заключение и вопросы**

Вот. Ну что, здесь интересного во всём этом? Значит, ребята, интересно, что создана модель, которая убедительно показывает, что есть закономерности, которые неизвестны в настоящее время науке. И больше того, есть ряд учёных, которые считают, что они, в принципе, не могут существовать эти закономерности. Мы же их обнаружили, эти закономерности, и исследовали, и описали в виде функций различных. Ну я так думаю, что это большое достижение, по моему мнению. Кто бы что ни говорил по этому поводу. А говорят так, что это вообще невозможно, в принципе, потому что невозможно никогда. А вот мы демонстрируем подробнейшим образом с исходными данными, с моделями, способами расчёта, что это возможно. И я это описываю в такой форме, ребята, которая выглядит как инструкция практически, инструкция, которую можно выполнить и самому лично во всём этом убедиться.  
Так что, ребята, я вам раскрыл вопрос, который, по-моему, раньше обошёл вниманием. А теперь мы, а теперь я вижу, что мы пришли к пункту обсуждения, на котором мы остановились в прошлый раз. Это пункт 1.11.3.  
Так, ребята. Ну что, у нас ещё осталось довольно много времени. Вот, но у меня есть сомнения, что надо это сейчас рассказывать. Я оставил бы на следующее занятие. Потому что это вопрос большой, самостоятельный, интересный, который я хотел бы оставить на следующее занятие. А сейчас давайте задавайте вопросы, какие у вас возникли, и мы их обсудим, я попробую на них ответить. Может быть, я надеюсь, у вас возникли вопросы из того, что я рассказывал вам. Или, может быть, у вас есть вопросы по вашим работам. Кстати, я хочу вам сообщить приятную новость, ребята. Послушайте сейчас внимательно. Вы слушаете, нет? Очень приятная новость. Рассказать? Конечно. Спрашивает, да? Значит, ну, в общем, какой-то студент захотел зарегистрироваться в какой-то системе, которая платная вообще эта система. Но, значит, можно для студента зарегистрироваться бесплатно, если он подтвердит свой статус студента с помощью корпоративного адреса электронной почты. И стали на этом форуме, на группе кафедральной, стали все преподаватели это обсуждать. И можем ли мы его загубить этого студента, не дать ему этой возможности, или можем дать? И кто, кто мы такие, если мы не дадим такой возможности? И в таком плане примерно. А что там Марк давал? Я так делал уже. Что ты делал? Ну я не знаю, что это такое. Brain - это что-то связано с мозгом, наверное, волны. А! Так это, может быть, ты был или не ты? Там не назывался студент. Короче говоря, ну, понятно. Ну ты не единственный, кто пытался войти в какую-то систему такую. Короче говоря, вопрос этот стали обсуждать. Ну и тут я так посмотрел, посмотрел на всё это, слушал, слушал это всё, и взял и написал, что вообще-то в других вузах, как студент поступает в ВУЗ, ему сразу даётся корпоративный адрес. Привёл конкретно название этих вузов, ну, в сокращённом виде. Сказал, что это само собой разумеющимся является, и что если этой почты нет, то это ущемляет права студентов на регистрацию в социальных научных сетях, системах. И стал вот это всё написал, короче, это вот. И Креймер Алексей Семёнович, наш директор, начальник ЦИТа, центра информационных технологий, написал, что пусть приходит ко мне. Дайте ему мой телефон, и пусть приходит ко мне. А я написал, что после этого, что хорошо было бы сделать это на постоянной основе, а не так, что там кому-то там дать телефон там, и только по вашему разрешению. А хорошо бы, чтобы это все студенты имели право такое прийти и получить эту почту. И к моему удивлению, он написал Алексей Семёнович, это меня несколько удивило, потому что раньше он вообще наотрез отказывался давать эту почту. Просто отказывался. Ну, как будто отдельно помогал по большой просьбе зав кафедры или моей просьбе, он помогал. Ну, по старой памяти, потому что он когда-то работал у нас доцентом на кафедре. Ну и сейчас работает доцентом, но основной директор центра информационных технологий. И вот вдруг он отвечает, представьте себе: пусть в организованном порядке они дают мне заявку, кому дать адрес электронной почты. Правда, я написал там, в каком там централизованном порядке, он, по-моему, не уточнил. Сейчас я посмотрю. Да, кстати, похоже, что один студент, поздравляю вас. Вот. Я написал, что надо бы, конечно, без запроса специального. Не все ж могут этот запрос оформить. И в какой форме, на кого этот запрос оформлять? Может быть, по электронной почте это можно или как само собой разумеющееся. И мне Алексей Семёнович не ответил и в группе не ответил на это по этому поводу.  
Есть какие-нибудь вопросы, ребята, у вас? Ещё 10 минут занятия. Давайте спрашивайте что-нибудь, кого что волнует.  
Ну всё, ребята, до свидания. Всего самого хорошего. Может быть, когда-нибудь в будущем обратно поставлю в меню их. Или изменю остальные вещи, которые я, ну вот там уже всего поизменял, вставлю в ту старую версию то, что я поизменял. До свидания. До свидания. До свидания. До свидания.