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

***Российская Федерация***

**46 Лабораторная работа № 2 по дисциплине "Интеллектуальные технологии и представление знаний". Вопрос - Лабораторная работа 3.03. 2020-11-19**

**Заголовок:** ЭИДОС: Синтез, верификация моделей и важность оценки их достоверности

**Резюме текста:**

Семинар посвящен лабораторной работе №2 по дисциплине "Интеллектуальные технологии представления знаний" для группы 413, проводимой профессором Луценко Е.В. Основная тема – синтез и верификация моделей с использованием автоматизированной системы когнитивного анализа ЭИДОС-Х++.

**1. Введение и настройка:**

* **Начальная настройка и контекст:** Обсуждается продолжение лабораторной работы 3.03, фокусируясь на разработке интеллектуальных приложений и переходя к теме синтеза и верификации моделей.
* **Проблемы со входом и устранение неполадок:** Студенты сталкиваются с трудностями при входе в систему на компьютерах в аудитории, так как не помнят свои учетные данные (логин/пароль/email). Профессор помогает, объясняя формат университетского email (S + номер студбилета + @kubsu.ru) и отправляет его в чат, подчеркивая важность знания своих данных для доступа.
* **Запуск системы ЭИДОС:** Даются пошаговые инструкции по запуску системы: найти папку Eidos на диске X в Program, выбрать подпапку с номером, соответствующим номеру студента в списке группы, и запустить файл start\_eidos.exe.
* **Технические трудности:** Возникают проблемы с зависанием системы при запуске у одного студента и задержкой демонстрации экрана в Teams по сравнению с реальными действиями. Профессор объясняет процесс обновления и распаковки файлов системы при первом запуске.

**2. Основная задача: Синтез и верификация моделей:**

* **Загрузка приложения:** Студентам показывают, как загрузить необходимое приложение для лабораторной работы ("Идентификация предметов по их признакам, Лаб. работа 3.03") из облака через интерфейс ЭИДОС (режим 1-3).
* **Экран импорта данных:** Кратко обсуждается окно настройки импорта данных из Excel, которое студенты должны помнить с прошлого занятия (диапазоны шкал, типы файлов).
* **Создание модели и обучающей выборки:** После подтверждения настроек импорта система автоматически создает классификационные и описательные шкалы, кодирует исходные данные и формирует обучающую выборку, готовую для синтеза моделей.

**3. Понимание моделей и мер ЭИДОС:**

* **Статистические модели (Abs, PRC2):** Рассматриваются матрицы абсолютных частот (Abs) и относительных частот (PRC2). Подчеркивается разница между абсолютными и относительными частотами для оценки характерности признака для класса. Пример с тремя встречами признака "размер под руку" в 3 объектах класса "средства связи" и 7 объектах класса "элементы компьютера" показывает, что признак более характерен для первого класса (100% против 43%).
* **Системно-когнитивные модели (Inf1, Inf3):** Объясняется смысл меры количества информации по Харкевичу (Inf1), показывающей информативность признака для принадлежности к классу. Рассматривается модель Хи-квадрат (Inf3), отмечается её сбалансированность (суммы по строкам/столбцам равны нулю), проводится аналогия с магическими квадратами.

**4. Оценка достоверности моделей (Ван Рисберген):**

* **Важность достоверности:** Профессор акцентирует внимание на критической важности проверки достоверности любой модели *перед* её использованием для выводов, прогнозов или принятия решений. Использование недостоверной модели может привести к ошибочным и даже катастрофическим последствиям.
* **Подход Ван Рисбергена:** Объясняется метод оценки достоверности моделей через решение задачи идентификации, предложенный Ван Рисбергеном. Вводятся понятия истинно положительных (TP), истинно отрицательных (TN), ложно положительных (FP) и ложно отрицательных (FN) решений.
* **Точность, полнота, F-мера:** Упоминаются основные метрики, рассчитываемые на основе сумматоров Ван Рисбергена: точность (precision), полнота (recall) и F-мера (гармоническое среднее точности и полноты) как итоговая мера достоверности.
* **Практическое применение в ЭИДОС:** Демонстрируется режим 3-4 для количественного измерения достоверности моделей. Обсуждается возможность ускорения расчетов с использованием ГПУ (хотя в данной аудитории недоступно) и опция проверки достоверности на случайной подвыборке для экономии времени при работе с большими данными.

**5. Заключение и дальнейшие шаги:**

* Семинар завершается на обсуждении мер достоверности. Следующее занятие будет посвящено продолжению рассмотрения этого вопроса.

**Детальная расшифровка текста:**

**1. Введение и настройка**

* **1.1. Начальная настройка и контекст**
  + (0:02) года, третья пара,
  + (0:04) которая идёт в 11:00,
  + (0:06) 13:00
  + (0:08) с группой 413.
  + (0:10) Лабораторная работа номер два,
  + (0:13) на которой мы изучаем,
  + (0:15) э-э, сейчас расскажу, что, по дисциплине интеллектуальные технологии представления знаний.
  + (0:21) Занятие ведёт профессор Луценко Евгений Венеаминович.
  + (0:25) Мы проходим лабораторную работу 3.03.
  + (0:28) Я подробнейшим образом всё рассказываю и объясняю, всё, что касается разработки интеллектуальных приложений.
  + (0:35) И сейчас мы уже, у нас было занятие в прошлый раз очень хорошее в классе компьютерном.
  + (0:41) И сейчас мы продолжаем это занятие, и у нас вопрос такой, с которого мы начинаем рассмотрение -
  + (0:46) это синтез и верификация моделей.
  + (0:49) И мы делаем следующим образом. Значит, у вас кто-то один
  + (0:54) э-э, включает телефон на громкую связь,
  + (0:58) кладёт рядом с вами.
  + (1:00) Значит, во-первых, покажите мне, пожалуйста, аудиторию, где вы сидите.
  + (1:04) Я хотел бы посмотреть, сколько вас там человек.
  + (1:06) Как вы там выглядите хоть.
  + (1:11) Вот. И э-э, кто-то, э-э,
  + (1:13) открывает...
  + (1:16) Все все садятся за компьютеры.
  + (1:18) Я ж я ж так понимаю, вы сейчас в компьютерной аудитории, да?
  + (1:24) Ребят, отвечайте.
  + (1:26) Да. Да. Компьютер.
  + (1:28) Значит, включаете компьютеры,
  + (1:31) заходите в папочку, э-э, на диск X,
  + (1:34) Program.
  + (1:37) И папочку система Eidos.
  + (1:41) Вот. И, значит, на...
  + (1:44) выбираете такую... там папочки есть 01, 02, 03, 40 там, до 40.
  + (1:50) Выбираете папочку с таким номером, какой у вас номер в списке группы.
  + (1:57) И запускаете файл Start Eidos.
  + (2:00) А мне сейчас демонстрируйте экран
  + (2:03) одного компьютера. Ну, допустим, староста там у вас сейчас, нет? С вами?
  + (2:07) Вот.
  + (2:08) Нет. Нет, не с нами.
  + (2:10) Ну тогда, тогда кто-нибудь из вас наиболее сообразительный.
* **1.2. Проблемы со входом и устранение неполадок**
  + (2:16) Вот у нас проблема. Просто мы сидим с телефонов, а с компьютера мы не помним, как зайти именно в свой кабинет.
  + (2:25) Вот. Поэтому проблемно, наверное, будет зайти с компьютера.
  + (2:30) Вы же в прошлый раз спокойно же вошли и работали.
  + (2:33) Просто староста, она с компьютера дома заходит. Вот. И как бы она постоянно вводит.
  + (2:40) А у нас проблема.
  + (2:43) Ну так что проблема? В чём? В том, что вы пароль не помните или что? Или имя имя, пароль?
  + (2:52) Ну вот данные для входа.
  + (2:55) Мы просто как один раз зашли, а с это, с компьютера, ну, получается, с чужого.
  + (3:00) Непонятно.
  + (3:02) Вы должны это знать.
  + (3:04) А кто вам это может сказать? Вот если я сейчас позову системного администратора, он подскажет вам, Геннадия?
  + (3:11) Он знает именно пароли ваши или нет?
  + (3:16) Кто вам? Ну вы должны их хорошо знать эти имена и пароли. Вы же ж не не только на моём занятии занимаетесь на компьютерах. Что ж это такое? Компьютеры стоят, а вы не можете зайти.
  + (3:26) Это очень странно.
  + (3:27) Просто мы с... мы заходим в какие-то из дома, какие-то здесь, но я даже не знаю. Нам просто нужна почта, я, например, не помню свою почту. Именно вот как она там находится.
  + (3:42) Как она там называется.
  + (3:44) Ну я могу вам сказать вашу почту.
  + (3:46) Со своего именно компьютера я могу зайти.
  + (3:50) Потому что там уже все пароли введены, получается, всё введено.
  + (35:4) А здесь как бы, я даже не знаю, смогу ли я с тимса.
  + (3:59) А у вас с собой есть компьютеры, что ли, ваши?
  + (4:02) Нет. Здесь мы находимся на таком, стационарном.
  + (4:11) Значит, ваша почта:
  + (4:13) буковка S,
  + (4:16) потом номер вашего
  + (4:19) Вот девочка сейчас попробует зайти. студенческого билета вашего.
  + (4:24) Вот. Потом дальше лягушечка kubsu.ru, причём без всяких точек, без ничего, kubsu.ru подряд.
  + (4:33) Ну я сейчас вам даже, сейчас, пожалуй, вырежу это.
  + (4:37) Сейчас я это сделаю.
  + (4:43) Вот. Пошлю вам, э-э,
  + (4:45) вашу почту. Вообще меня поражает, что вы не знаете почту.
  + (4:50) Значит, вот, э-э, она написана на вашем этом
  + (4:54) окошке, беленьком вот этом. Я вот сейчас в чат послал.
  + (4:59) Написана почта.
  + (5:02) Что здесь не знать?
  + (5:08) Вот.
  + (5:12) Увидели, нет?
  + (5:17) Я в чат послал почту.
  + (5:21) Такая же точно почта у остальных, только, э-э, номер билета другой.
  + (5:30) Так звонить звонить администратору или не надо?
  + (5:35) Нет, не надо. Девочка вот попробует зайти сейчас.
* **1.3. Запуск системы ЭИДОС**
  + (5:45) А ещё я могу позвонить замдекана, но он тоже вам подскажет, как войти.
  + (5:51) Ну я этого не советую делать,
  + (5:54) доводить до этого.
  + (6:05) Вот. Значит, на диске X, там папочка Program есть, в ней Eidos. И выбираете там папочку, в папочке Eidos есть 1, 2, 3, там они подряд идут, соответствующую вашему номеру в списке группы.
  + (6:18) Наверное ж вы все знаете номера свои, да, в списке группы?
  + (6:21) О! И запускаете Start Eidos файлик.
  + (6:26) Я увидел. Молодцы.
  + (6:28) Start Eidos, он самый в начале находится, когда по алфавиту рассортировано.
  + (6:32) Сейчас как раз по алфавиту.
  + (6:34) То есть запускаете файл Start Eidos X.
* **1.4. Технические трудности**
  + (6:42) А что вы делаете?
  + (6:53) Я не понял, почему вы не запускаете систему?
  + (6:57) Вот девочка открыла, у неё всё зависло.
  + (7:05) Мм. Ну вот на Start Eidos нажимаете и Enter или два раза. Запустить.
  + (7:11) Значит, сейчас скачается обновление, возможно.
  + (7:19) И запустится архиватор на разархивирование.
  + (7:35) Вот, вот архиватор запустился. Надо нажать...
  + (7:38) Где это окошко архиватора? Оно там на панели задач должно быть где-то.
  + (7:49) Демонстрация экрана явно отстаёт от того, что есть.
  + (7:54) Ну, у вас, в общем, окошко...
  + (7:56) Мы видим одно, а на экране в Тимсе другое.
  + (8:00) Понятно. Ну, весь университет сидит, наверное ж.
  + (8:04) Значит, окошко архиватора должно открыться, раровского. Надо развернуть, заменить все файлы.
  + (8:10) Прямо вот там, в той же папочке запустить. Э-э, уже, по-моему, один раз вы это делали, если не ошибаюсь.
  + (8:18) Вот.
  + (8:21) Появится сейчас, э-э, скачивание произойдёт, потом архиватор появится.
  + (8:28) И нужно там не отменять, а разархивировать, заменяя все файлы.
  + (8:40) Ну у нас получилось вот что. Давайте я попробую сейчас...
  + (8:47) Ну один... Это вы хотите показать по телефончику?
  + (8:50) Вот, да. Получается, вот то, что мы видим и то, что на экране совсем другое.
  + (8:56) Это всё правильно.
  + (8:58) И что нам нужно?
  + (8:58) На развернуть
  + (9:01) и когда нас спросят все, для всех файлов.
  + (9:04) Да, для всех.
  + (9:07) Так же сделайте на других, на всех компьютерах,
  + (9:11) Угу.
  + (9:13) где вы работаете. И желательно, чтобы вы каждый работали на своём компьютере.
  + (9:18) Он выдал "Прекращена работа программы".
  + (9:26) Какой?
  + (9:34) Ну мне нужно видеть экран того... Если есть какие-то проблемы, то тот компьютер, на котором проблемы, мне нужно видеть экран его.
  + (10:02) Как можно поделиться экраном? Просто вот мы делимся, и он делится папкой.
  + (10:10) А сам экран, он непонятно, как сделать.
  + (10:14) Вот рабочий стол...
  + (10:15) Когда вы нажимаете вот "Поделиться экраном", вот тут нужно выбирать самое левое окошко там.
  + (10:22) Там появляются много окошек: текущее окошко, активное... И самое левое - это экран.
  + (10:29) Вот нужно выбирать самое левое окошко. Тогда весь экран будет, э-э, делиться.
  + (10:35) Нельзя поделиться всем экраном.
  + (10:38) А здесь нельзя поделиться всем экраном.
  + (10:43) Очень странно.
  + (10:46) В прошлый раз тоже поделилась всем.
  + (10:48) Сейчас вы поделились самой самим экраном, где система Eidos. Ну у вас уже новая версия стоит, всё нормально.
  + (10:54) Вот так же должна быть 17.11 у всех остальных.

**2. Основная задача: Синтез и верификация моделей**

* **2.1. Загрузка приложения**
  + (11:03) Вот. Ну будем, э-э, так считать, что другие сейчас тоже это же самое сделают всё.
  + (11:09) Вот. Ну, а попробуйте открыть режим 1-3.
  + (11:19) Сейчас только что я занимался с другой группой. Какая у вас аудитория сейчас, ребят?
  + (11:26) 202N.
  + (11:28) Вот в этой же группе занимались сейчас только что предыдущая группа, то есть в этой же аудитории 412 группа занималась.
  + (11:35) У них совершенно спокойно весь экран они демонстрировали.
  + (11:40) Там видел рабочий стол, там с иконками, всё нормально было.
  + (11:44) Это самое левое окошко, когда делитесь экраном, самое левое.
  + (11:49) Скрин 1 называется.
  + (11:52) Не собрание там, не Total, ничего, а вот именно Скрин 1.
  + (11:58) И прекрасно всё видно.
  + (12:13) Так, ну и давайте теперь запускайте режим 1-3.
  + (12:21) Вот. Тоже мы видим
  + (12:25) вполне.
  + (12:41) Ещё раз покажите экран, пожалуйста.
  + (12:46) Только весь уже.
  + (12:52) Там когда я вам объяснял, как сделать показ экрана, то, э-э, и вот это моё такое изображение, что ли?
  + (13:00) У нас получилось.
  + (13:02) Вот. Ну да, самое левое окошко, скрин 1, оно и есть.
  + (13:07) Значит, здесь теперь нажимаем режим, э-э, кнопочку нажимаем "Скачать приложение из облака".
  + (13:14) Вот там, где "Добавить лабораторную работу", там есть "Скачать приложение из облака".
  + (13:24) Вот жирненьким шрифтом внизу написано, а это правее кнопочка "Скачать приложение из облака".
  + (13:31) И в самый-самый верх этого списка двигаетесь.
  + (13:34) Вообще в самый верх.
  + (13:36) Ну можно бегунком, можно Ctrl+Home, Ctrl+Page Up.
  + (13:44) Вот, это облачные Eidos-приложения, которые сейчас на данный момент есть. И выбираем здесь, э-э,
  + (13:50) идентификация предметов по их признакам, лабораторная работа 3.03.
  + (13:56) Ставим курсор на него
  + (13:58) и нажимаем "Установка Eidos приложения".
  + (14:01) Внизу кнопочка жирным шрифтом выделена.
  + (14:05) Появляется каталог в стиле Юникса, который показывает, что там у нас есть в этой папочке на FTP-сервере системы.
  + (14:13) Закрываем это окошко. Это просто предупреждение, что будет скачиваться.
  + (14:17) Начинают скачиваться вот эти файлы.
* **2.2. Экран импорта данных**
  + (14:24) И сообщение, что там есть файлы, которые похоже, что какое-то описание содержат. О'кей нажимаем.
  + (14:30) Появляется вот такое окошко. Значит, сейчас, ребята, вы должны вспомнить. Я вам рассказывал про про это окошко или нет?
  + (14:36) Это э-э, экранная форма программного интерфейса ввода данных из таблиц экселевских в систему.
  + (14:42) Там есть вот диапазон классификационных шкал со второй по третью, диапазон описательных шкал с четвёртой по двенадцатую, равные интервалы, тип файла старый Excel.
  + (14:54) Это я всё это должен был вам рассказывать на прошлом занятии, судя по всему.
  + (14:58) Судя по тому, что я записал себе.
  + (15:04) Ребят, вы были на прошлом занятии, нет?
  + (15:08) Да, были.
  + (15:09) Вы это помните, нет, чтобы я это объяснил вам?
  + (15:13) Да, да, помню.
  + (15:15) Слава Богу.
  + (15:16) Значит, вы должны сразу реагировать, без таких диких пауз длительных.
  + (15:20) Вот. Хотя я понимаю, что тут, может быть, это не вы паузу делаете, а там Teams или ещё кто-нибудь, не знаю, там.
  + (15:27) Через Америку преподаём, не знаю. Значит, О'кей нажимаем.
* **2.3. Создание модели и обучающей выборки**
  + (15:37) Запускается Excel, потому что там конвертер экселевский использовался для преобразования экселевского файла в DBF.
  + (15:45) И вот здесь вот у нас экранная форма, на которой показано, что мы нашли две классификационных текстовых шкалы и одну, э-э, числовую описательную, и ещё несколько текстовых шкал.
  + (16:01) И вот это, поскольку есть числовая шкала, то спрашивает система, сколько числовых диапазонов брать.
  + (16:06) Пусть будет всё по умолчанию.
  + (16:08) Но если вы поменяли, допустим, захотели бы поменять, то надо потом пересчитать шкалы и градации.
  + (16:15) И вверху здесь выводится, сколько у нас будет колоночек в таблице. Видите, 14 на 50.
  + (16:20) Вот в этой экранной форме написано, что у нас будет 14 колоночек классов и 50 колоночек признаков.
  + (16:26) Сейчас выходим на создание модели просто.
  + (16:33) Вот, смотрите, что произошло. Созданы классификационные описательные шкалы и градации, и они использованы для кодирования исходных данных.
  + (16:40) Получилось в результате обучающая выборка. О'кей нажимаем.
  + (16:47) Выполняем теперь вот это, что здесь написано. Вы прочитали, что там написано, ребята?
  + (16:54) Не успели?
  + (16:55) Значит, давайте...
  + (16:57) У меня не было.
  + (16:57) Ну я так и понимаю. Вот студенты почему-то всегда там нажимают О'кей сразу.
  + (17:02) Надо такими большими буквами красными писать: "Не нажимайте О'кей, прочитайте, что написано на окошке".
  + (17:08) Вот. Закрывайте...
  + (17:08) Надо сначала написать "Прочитать", а потом нажимать О'кей.
  + (17:11) Ну конечно, конечно. А студенты сразу нажимают О'кей, и потом я рассказываю вот это, что я буду, то, что там написано, рассказываю.
  + (17:20) Значит, закрывайте это окошко с жёлтым фоном.
  + (17:26) Вот. И, э-э, видите, у нас уже установлена лабораторная работа 3.03. Закрываем это окошко
  + (17:33) диспетчера приложений.
  + (17:35) И смотрим на, э-э,
  + (17:38) сам файлик исходных данных.
  + (17:40) Но, по-моему, всё это я на прошлом занятии рассказывал, поэтому сейчас нам это делать даже не нужно, смотреть на файл исходных данных.
  + (17:47) Эселевский файл. Помните, я показывал там мышки там,
  + (17:50) клавиатуры, экраны.
  + (17:53) Да, да, да.
  + (17:54) Вот.
  + (17:55) Значит, мы дошли до того, что нам нужно теперь создать модели и уже дальше изучать, что там за модели, какая у них достоверность и так далее.

**3. Понимание моделей и мер ЭИДОС**

* **3.1. Статистические модели (Abs, PRC2)**
  + (18:03) Модели создаются в режиме 3-5. Да, кстати, подождите, зайдите в режим 1-3.
  + (18:11) И нажмите помощь там слева кнопочка.
  + (18:16) Здесь по пунктикам написано то, что я сейчас вам говорю делать.
  + (18:21) Вот сейчас мы будем смотреть режим 3-5. Остальное мы сделали на прошлом занятии.
  + (18:26) Закрываем это.
  + (18:28) И смотрим на режим 6-4. Это тоже закрываем.
  + (18:34) 6-4.
  + (18:36) И разворачиваем по ширине на всю ширину,
  + (18:40) на на весь экран это.
  + (18:42) Разворачиваем и на ширину, там, где вот у нас 101%, там справа есть по ширине листа.
  + (18:47) Нет, справа, справа от этого процента. Ну можно так.
  + (18:51) На плюсик нажимать. Это не это весь лист вы смотрите. А надо там чуть левее нажать, тогда будет на ширину листа.
  + (18:58) Вот это, да.
  + (19:01) Вот. Значит, что мы делаем в системе Eidos? Мы делаем всего лишь несколько вещей. Мы анализируем исходные данные,
  + (19:08) находим там прошлые и будущие события,
  + (19:11) э-э, создаём соответствующие справочники прошлых и будущих событий и кодируем исходные данные с помощью этих
  + (19:17) справочников, получаем обучающую выборку,
  + (19:21) и потом, отражающие события, которые моделируются в этой модели, описываются этими данными.
  + (19:26) И потом находим причинно-следственные связи между этими событиями. То, что сейчас мы будем делать в режиме 3-5.
  + (19:33) И получается уже модель.
  + (19:39) Информация о силе и направлении причинно-следственных связей между значениями факторов и переходом объекта управления моделирования в те или иные будущие состояния, как желательные, так и нежелательные.
  + (19:51) Потом мы эту информацию можем применить для решения различных задач:
  + (19:55) для задач идентификации, прогнозирования, принятия решений,
  + (19:58) исследования моделируемой предметной области.
  + (20:01) Когда мы применяем эту информацию для решения задач управления, то есть достижения цели, то она превращается в знание.
  + (20:08) Следующую страничку смотрите
  + (20:11) и, э-э,
  + (20:13) там тоже по ширине выравниваете изображение.
  + (20:21) Ну курсор надо поставить сюда на страничке и потом нажать эту кнопочку.
  + (20:28) Вот.
  + (20:29) Что-то как-то не особо оно, э-э, ну по ширине надо сделать, иначе не видно ничего.
  + (20:37) Ну это всё, так не получается.
  + (20:39) Тогда плюсик нажимаете, тогда плюсик нажимаете.
  + (20:41) Или там 200 укажите. Плюсик нажимаете, будем смотреть, что получится.
  + (20:47) Ага. Ну тогда надо прокрутить вправо, наверное, да? Так.
  + (20:52) Ну ладно, в принципе видно.
  + (20:54) Значит, у нас первый пункт с сереньким фоном, видите?
  + (20:57) Когнитивно-целевая структуризация предметной области.
  + (21:00) Это единственный не автоматизированный этап
  + (21:02) системно-когнитивного анализа.
  + (21:05) Не автоматизированный в системе Eidos.
  + (21:07) На этом этапе мы определяем, что у нас в качестве факторов будет рассматриваться или признаков,
  + (21:11) из значения свойств.
  + (21:14) А что в качестве результатов их влияния этих факторов, какие, э-э, будущие состояния будет объект принимать под их действием.
  + (21:23) И вот эти, э-э, состояния классифицируются с помощью классификационных шкал,
  + (21:28) а факторы с помощью описательных шкал.
  + (21:31) То есть система Eidos содержит подсистему, которая так и называется "Формализация предметной области",
  + (21:36) которая обеспечивает анализ исходных данных, формирование справочников
  + (21:40) э-э, классификационных и описательных шкал и градаций, то есть классов вот этих состояний объекта и факторов, действующих на него, или значения свойств.
  + (21:49) С их помощью кодируются исходные данные, получается обучающая выборка.
  + (21:54) Потом следующий режим - это синтез и верификация модели. Прокрутите чуть-чуть вниз.
  + (22:00) На этом в этом режиме, это режим 3-5, мы создадим модели статистические: матрицу абсолютных частот,
  + (22:06) матрицу условных и безусловных процентных распределений.
  + (22:11) Потом с их помощью рассчитываются матрицы системно-когнитивных моделей:
  + (22:15) Inf1-Inf7.
  + (22:17) Выбирается наиболее достоверная модель по её критерию Ван Рисбергена, и в этой наиболее достоверной модели решаются задачи
  + (22:24) распознавания, идентификации, прогнозирования,
  + (22:27) принятия решений
  + (22:29) или управления и исследования моделируемой предметной области.
  + (22:33) Здесь эти задачи показаны в виде отдельных блоков,
  + (22:36) но они между собой тесно взаимосвязаны.
  + (22:39) Мы эту взаимосвязь рассмотрим позже.
  + (22:41) А сейчас только могу сказать, что это вот разделение, оно несколько условно. На самом деле эти задачи переплетаются.
  + (22:49) Вот вы можете чуть-чуть прокрутить вниз и увидите там, что я написал там, как они соотносятся друг с другом,
  + (22:55) как они взаимосвязаны.
  + (22:57) Но сейчас это не будем мы это рассматривать, сейчас это у нас такой задачи нет, но там это написано.
  + (23:03) Закрываем эту форму.
  + (23:05) Это просто я вам коротко описал, чем мы сейчас будем заниматься.
  + (23:11) Вот вы увидели краткое описание.
  + (23:14) Оно краткое, но содержательное, конкретное описание,
  + (23:17) как они соотносятся эти задачи.
  + (23:20) Дальше мы закрываем эту этот Help
  + (23:24) 6-4
  + (23:25) и переходим в режим 3-5.
* **3.2. Системно-когнитивные модели (Inf1, Inf3/Chi-squared)**
  + (23:30) Я где-то на минуту позже всё говорю.
  + (23:35) То есть есть задержка по времени хорошая.
  + (23:40) Ну, широкополосный интернет.
  + (23:43) Режим 3-5. Запускаем его и смотрим. Что мы здесь видим?
  + (23:48) Мы видим название модели статистических
  + (23:51) и системно-когнитивных.
  + (23:54) Видим, что мы можем вести расчёты на центральном процессоре и на графическом процессоре.
  + (23:58) Значит, на ваших компьютерах вот в университете на графических процессорах не получается.
  + (24:03) То есть там видеокарта не Nvidia,
  + (24:06) не фирма Nvidia и не не другой фирмы, которая бы использовала чипсет Nvidia.
  + (24:11) То есть не GeForce, например.
  + (24:14) То есть будем использовать центральный процессор. Поставьте на ЦПУ, вот там внизу справа есть ЦПУ,
  + (24:19) графический процессор и ЦПУ.
  + (24:23) Вот поставьте птичку на ЦПУ переключатель.
  + (24:27) Значит, могу вам сказать, что на графическом процессоре можно считать те же самые задачи, что на центральном, примерно там, ну, от сотен до тысяч раз быстрее.
  + (24:38) Я наблюдал до 4.000 раз быстрее расчёты на графическом процессоре, но очень такой простенькой видеокарте, которая на моём компьютере GT 240.
  + (24:48) Там всего всего лишь навсего 96 шедерских процессоров.
  + (24:52) А на хороших видеокартах, которые правда довольно дорогие, Nvidia 1700, то есть даже есть и выше уже, они могут стоить дороже компьютера.
  + (25:00) Там уже около 4.000 шедерских процессоров.
  + (25:04) Каждый из них мощнее центрального процессора компьютера.
  + (25:08) Хотя в центральном есть ядра вычислительные, вот у меня, допустим, i7 процессор, там четыре ядра физических и четыре логических в моём процессоре именно конкретно.
  + (25:18) Вот. А представьте себе, что на маленькой простенькой видеокарте 96 ядер вычислительных,
  + (25:24) физических.
  + (25:25) И каждый из них по э-э, мощи вычислительной примерно как процессор центральный.
  + (25:30) То есть это уже в сотни раз ускорение расчётов на некоторых задачах.
  + (25:35) А я наблюдал до 4.000 раз. Ну это зависит от того, какая размерность модели, сколько объектов обучающей выборки и тому подобное.
  + (25:43) Значит, теперь смотрите, ребята, значит, у нас может быть очень много объектов обучающей выборки.
  + (25:48) Мы проверяем достоверность модели путём решения задачи идентификации.
  + (25:53) То есть мы эти объекты обучающей выборки копируем в распознаваемую выборку. Когда вот стоит птичка вот там, где у нас написано "Параметры копирования обучающей выборки в распознаваемую", видите?
  + (26:04) Всю выборку стоит сейчас по умолчанию.
  + (26:07) Так вот, если, допустим, у нас там не 20 было бы объектов, а 4 млн, к примеру,
  + (26:12) и было бы, допустим, 1.500 классов, а каждый, э-э, класс описывался бы, допустим, там 30.000 признаков.
  + (26:20) Это всё возможно, ребята, в системе Eidos.
  + (26:23) Вот тогда представьте себе, у нас каждый объект, сначала был бы цикл по объектам выборки распознаваемой, да, тестовой.
  + (26:30) Потом бы для каждого объекта был бы цикл по всем классам от одного до 1.500.
  + (26:35) И чтобы сравнить объект с каждым из классов, было бы 30.000 сумм произведений.
  + (26:39) То есть суммирование э-э, 30.000 произведений.
  + (26:44) Ну, вы представляете себе, что это такое? Это чудовищное количество расчётов.
  + (26:48) И вот это могло бы затянуться на полгода, к примеру бы.
  + (26:52) Поэтому я сделал такую возможность, э-э, только синтез модели осуществлять или только верификацию.
  + (26:57) Это было ещё до использования графического процессора сделано.
  + (27:02) Значит, сейчас какая есть возможность? Мы можем синтез модели даже на очень больших данных осуществлять всё в разумные сроки.
  + (27:08) Вот, скажем, если там было 4 млн
  + (27:11) э-э, объектов, ну, допустим, там, э-э, несколько часов она поработала бы, допустим, создала бы эти модели,
  + (27:19) не не модели, а шкалы, шкалы и обучающую выборку.
  + (27:24) Вот, а потом синтез модели. Синтез он мог бы продолжаться там, скажем, минут 15-20 вот таких на такой модели.
  + (27:31) И всё, понимаете? То есть это вполне разумные реальные сроки.
  + (27:35) Потом мы могли бы взять вот этот режим, смотрите, где слева написано, не копирование, написано "Копирование N случайных объектов".
  + (27:42) Вот поставьте туда птичку.
  + (27:46) Вы тогда видите, что появляется окошко, где можно ввести 100, например, или 1.000.
  + (27:51) И запустить только верификацию после уже синтеза модели.
  + (27:56) Видите, да? Птичку справа поставить "Только верификация", а в окошке ввести 1.000.
  + (28:01) Тогда бы у нас, допустим, ну, к примеру, я просто к примеру говорю, 8 минут там считается. Система потом показывает время с точностью там до секунд, но считает она его с э-э, сотыми долями секунды это время измеряет.
  + (28:18) И вы можете взять, построить пропорцию.
  + (28:21) Допустим, 1.000 объектов считалось верификация, э-э, 8 минут.
  + (28:28) А X объектов будет считаться 480 минут.
  + (28:33) Ну, то есть 8 часов.
  + (28:36) То есть вы предполагаете, что вы вот так вот, э-э, измерили, с какой скоростью она это делает, и решили, что давай-ка, пожалуй, я пойду спать, а утром пусть он уже будет результат получен.
  + (28:46) И я посмотрю, какая достоверность модели.
  + (28:49) И вот, считая эту пропорцию, вы получаете, что для этого, чтобы посчитать за 8 часов достоверность модели, вам нужно в этом окошке задать 48.000 там 345, допустим, этих объектов выборки. Вы задаёте там округлённо, ну, допустим, 50.000, чтобы легче было потом рассказывать.
  + (29:07) И идёте спать. Утром просыпаетесь, там осталось там немножко совсем, раз досчиталось за 9 часов.
  + (29:13) Всё, вы получаете, э-э, возможность узнать, сколько времени займёт тот или иной процесс,
  + (29:19) таким вот образом, сами его прогнозируя.
  + (29:22) Теперь смотрите внизу, вот тут есть информация интересная, что вот эта модель составляет собой шесть десятитысячных процента от того, что может обрабатывать система Eidos.
  + (29:32) То есть это значит, что вот у нас, э-э, это число объектов выборки, это размерность модели, она совершенно миниатюрная для системы Eidos. То есть её возможности колоссальные в этом отношении, она может просто огромные модели рассчитывать.
  + (29:48) Э-э, значит, я пробовал использовать функции системы, которые в ней реализованы для того, чтобы создать матрицы 100.000 на 100.000, например.
  + (29:55) А это матрица 14 на 50.
  + (29:59) Я имею в виду 100.000 колонок на 100.000 строк.
  + (30:02) То есть вы понимаете, что это совершенно несопоставимо с этой вот маленькой моделью.
  + (30:07) Значит, но система определяет, может ли она посчитать такую модель, которую там у нас сейчас, э-э, мы хотим посчитать её
  + (30:15) при таком числе классов, в таком числе признаков, в таком числе объектов выборки. И определяет, что, допустим, она это может сделать, вот как в данном случае, шесть десятитысячных процента.
  + (30:26) А если определяет, что сделать не может,
  + (30:31) потому что не хватит оперативной памяти или места на диске, то она даже не будет делать синтез, она сделает только оценку достоверности, идентификацию, распознавание.
  + (30:41) Вот в этом окошке, где сейчас написано 100, можно написать пять, например.
  + (30:45) Запускаем. Система очень с высокой достоверностью определяет степень достоверности своих решений.
  + (30:52) Извините за каламбурчик. То есть она вот решает задачу такую... Да, поставьте птичечку, пожалуйста, "Копировать всю обучающую выборку".
  + (31:00) Вот так вот.
  + (31:01) То есть она решает задачу сравнения объектов с классами и оценивает степень достоверности этого сравнения. То есть она говорит: "Этот объект относится к этому классу, и уровень достоверности этого решения 1%".
  + (31:16) Ну это означает, что она очень сильно сомневается, что он к нему относится. Но всё-таки считается, что относится, потому что всё-таки уровень сходства положительный.
  + (31:24) Поэтому, значит, решение положительное.
  + (31:27) И такие решения с таким уровнем сходства обычно бывают ложными. То есть это ложное срабатывание.
  + (31:33) Система это прекрасно знает и сообщает, что очень низкая достоверность.
  + (31:37) Вот её оценка достоверности, я в чём, значит, я очень много раз в этом убеждался, и есть формы выходные, которые это очень убедительно, достоверно демонстрируют.
  + (31:47) Они очень адекватны. То есть она очень хорошо оценивает достоверность своих решений. Эта оценка, которую даёт система, адекватная.
  + (31:55) Выходим теперь, нажимаем О'кей.
  + (31:58) Происходит процесс сначала синтеза моделей
  + (32:02) вот по очереди, а потом верификация этих моделей, проверка их на достоверность.
  + (32:42) Вот, процесс завершён.
  + (32:45) Все модели проверены на достоверность. Теперь нажимаем О'кей.
  + (32:49) Созданы и проверены на достоверность.
  + (32:51) Теперь нас интересует, э-э, вопрос о том, а что ж там за модели созданы?
  + (32:56) Надо бы на них посмотреть.
  + (32:58) Давайте запускайте режим 5-5 и смотрим на эти модели.
  + (33:07) Вот, мы видим здесь слева три модели статистические:
  + (33:11) Abs, PRC1 и PRC2. Это
  + (33:14) матрица абсолютных частот и матрица условных и безусловных процентных распределений,
  + (33:19) рассчитанные двумя способами.
  + (33:22) А дальше идут системно-когнитивные модели: Inf1, Inf2 и так далее.
  + (33:28) Значит, открываем матрицу Abs.
  + (33:37) Что мы здесь видим?
  + (33:38) Мы видим, э-э, колонки - это классы, это градации классификационных шкал.
  + (33:44) А строки - это у нас, э-э, шкалы описательные и их значения.
  + (33:48) Когда шкала текстовая, то значениями являются уникальные текстовые значения в этой колонке, соответствующей шкале. То есть колонка - это шкала в исходных данных,
  + (33:58) вот, либо классификационная, либо описательная.
  + (34:01) Если шкала текстовая, то значениями являются уникальные текстовые значения в этой колоночке.
  + (34:08) Если числовая, прокрутите вниз, пожалуйста, немножко.
  + (34:14) То там есть диапазоны. Видите, размер два.
  + (34:18) Первый из двенадцати, двенадцатый из двенадцати и конкретно, в чём он заключается этот диапазон, от какой величины до какой.
  + (34:26) Что это такое, ребята? Это у нас найдено минимальное и максимальное значение
  + (34:30) в этой колонке
  + (34:31) размер два файла исходных данных, вот этого экселевского.
  + (34:37) И, э-э, разделено на то число диапазонов числовых, которое мы задали
  + (34:42) э-э, при вводе этой информации в систему.
  + (34:45) И система это может сделать двумя способами. Она может, э-э, либо одинаковые интервалы
  + (34:53) использовать числовые,
  + (34:57) как вот мы сейчас в этом случае задали, либо адаптивные,
  + (35:00) разного размера, с примерно одинаковым числом наблюдений.
  + (35:05) Э-э, почему примерно? Потому что наблюдение всегда целое число.
  + (35:08) Не бывает там половина наблюдения.
  + (35:12) Как вот не бывает полтора землекопа.
  + (35:15) Вот если три ямы роют два землекопа, то есть если две ямы роют три землекопа, то сколько одну яму роют?
  + (35:20) Полтора получается, да? 3 / 2.
  + (35:23) Ну все ж понимают, что одну яму роет один землекоп, а другую два,
  + (35:28) что так не бывает, чтобы полтора землекопа копали яму.
  + (35:32) Ну так же точно и здесь. То есть у нас число наблюдений всегда целое.
  + (35:36) И на диапазон будет приходиться число наблюдений, которое отличается на единицу,
  + (35:40) если адаптивные интервалы заданы.
  + (35:44) Они, откровенно говоря, конечно, получше, чем вот эти вот равные интервалы. Ну сейчас я простейший вариант вам показываю.
  + (35:50) Значит, э-э, в клеточках этой матрицы
  + (35:53) показано, сколько раз встретилось, э-э, определённое значение свойства у объектов определённой
  + (35:59) э-э, категории, класса.
  + (36:02) И там у нас, видите, единички, двоечки стоят, троечки кое-где.
  + (36:06) Если мы вправо сдвинемся, то мы увидим там и побольше числа.
  + (36:11) Вот. Почему? Потому что дальше у нас там идут колоночки, соответствующие обобщающим классам,
  + (36:17) и там больше примеров.
  + (36:19) Если мы сейчас двинемся вниз
  + (36:21) до конца,
  + (36:23) то мы увидим, что в системе есть информация о том, сколько объектов выборки обучающей было представлено по каждому из классов. И там у нас, видите, и один есть, и три,
  + (36:34) и четыре, и есть семь
  + (36:38) для класса "Элемент компьютера".
  + (36:43) Теперь поднимаемся до двадцать третьей строки.
  + (36:49) И мы здесь видим вот что: размер под руку, видите, свойство?
  + (36:52) И вправо двигаемся. Ну мы уже мы вот так уже вправо сдвинулись.
  + (36:57) Вот. Мы там видим, э-э,
  + (36:59) что это свойство "размер под руку" три раза встречается у средств связи
  + (37:04) и три раза встречается у элементов компьютера. Видите, да?
  + (37:08) 3 3 в двадцать третьей строке, да.
  + (37:11) Вот. И, э-э, возникает такой вопрос, который я хочу вам задать вот сейчас. Э-э,
  + (37:19) Пожалуйста, покажите мне аудиторию телефончиком.
  + (37:24) Я хочу посмотреть на на то, сколько там вас сидит и как вы выглядите.
  + (37:36) Секунду.
  + (37:38) О, я вижу клавиатуру.
  + (37:40) Ну у нас...
  + (37:42) Так, секунду.
  + (37:44) Не так много.
  + (37:45) Ну да.
  + (37:46) Сколько?
  + (37:48) Э-э, четверо.
  + (37:51) Понятно.
  + (37:54) Понятно.
  + (37:55) Ну что, ладно. Давайте дальше.
  + (37:58) Значит, вот хотелось бы, чтобы вы ответили на такой вопрос.
  + (38:01) Вот смотрите, значит, э-э,
  + (38:04) у нас этот признак три раза встретился у объектов,
  + (38:07) которые относятся к классу "Элемент компьютера", и три раза встретился у средств связи.
  + (38:12) И там, и там три раза он встретился.
  + (38:16) Видите?
  + (38:16) Можно ли из этого сделать такой вывод, что этот признак одинаково характерен и для средств связи, и для элементов компьютера?
  + (38:25) Или нет? Или у нас нет оснований такой вывод делать?
  + (38:34) То есть у этих объектов этих категорий этот объект встретился, этот признак встретился по три раза.
  + (38:40) То есть и у средств связи он три раза встретился, и у элементов компьютера три раза встретился.
  + (38:44) Можно ли сказать, что это означает, что этот признак одинаково характерен для этих категорий?
  + (38:50) Или нет? Нет оснований такой вывод делать?
  + (38:57) По идее характерен.
  + (38:59) Нет, он характерен, но в одинаковой степени или разной для этих классов?
  + (39:05) Ну то одинаково.
  + (39:07) Вот. Крутим вниз, ребята, вниз.
  + (39:12) В самый низ таблицы.
  + (39:15) Вот смотрите. Значит, средств связи у нас три было предъявлено объекта обучающей выборки.
  + (39:22) А элементов компьютера - семь.
  + (39:25) То есть получается, что у нас 100% средств связи с этим признаком, да, под руку, размер под руку. И где-то 43% э-э, элементов компьютера с этим признаком.
  + (39:36) То есть нельзя сделать вывод, на вот несмотря на эту строчку, если мы не знаем, сколько у нас объектов выборки по каждому из классов, то мы не можем сделать вывод о том, насколько этот признак характерен или не характерен для того или иного объекта, того или иного класса.
  + (39:53) Недостаточно у нас информации. То есть я вам спросил у вас, можно ли сказать, что он одинаково характерен?
  + (39:58) Нет, сказать так нельзя. Потому что мы ж не знаем, сколько объектов было предъявлено по классу средств связи и по классу "Элемент компьютера".
  + (40:08) Если одинаковое, тогда можно такой вывод сделать, что одинаково характерно. А если разное? Мы ж не знаем, одинаковое или разное.
  + (40:16) Вот. Вообще говоря, оно разное, конечно, число объектов.
  + (40:19) Получается, что мы не можем на основе этой матрицы делать какие-то обоснованные выводы непосредственно вот на основе этих вот абсолютных частот. А нам нужно перейти к относительным частотам, ребята.
  + (40:30) Что и делается в матрице PRC2. Закрываем эту форму.
  + (40:36) Смотрим матрицу PRC2.
  + (40:40) А здесь у нас, и тоже двигаемся вправо и на двадцать третью строчку.
  + (40:47) И мы здесь видим, что 100% э-э, средств связи и 42,8% элементов компьютера этим признаком обладают.
  + (40:58) То есть он для средств связи намного более характерен, чем для элементов компьютера.
  + (41:03) А вообще он характерен для этих классов или нет? Этот признак? Характерен, потому что в среднем по всей выборке всего 30% объектов этим признаком обладают.
  + (41:13) Посмотрите, пожалуйста, по всей выборке 30%. Классов 100, э-э, средств связи 100%.
  + (41:20) Ну это надо в правой колоночке смотреть безусловную вероятность. Вот давайте двадцать третью строку
  + (41:26) и в не и вправо до конца.
  + (41:30) Вот вы видите, значит, здесь у нас по всей выборке 30%, безусловная вероятность. У вас там курсор стоит сейчас.
  + (41:37) А, это у меня стоит курсор.
  + (41:39) Вот. А, значит, э-э, да, вот это.
  + (41:41) 30%.
  + (41:44) То есть получается так, что этот признак чаще встречается в этих категориях, чем в среднем по выборке.
  + (41:53) Ну, приведу вам такой пример. Значит, причём в одной категории он намного чаще встречается, в другой чаще, но не намного, чем по всей выборке. Ну сравнительно. Вот представьте себе, что нам нужно отличить студентов от студенток.
  + (42:06) И мы имеем такую же вот табличку, где у нас посчитано, какой процент студенток имеет длинные волосы, какой процент студентов имеет длинные волосы.
  + (42:16) И то же самое касается брюк. У кого брюки там есть у студентов, у студенток. И то же самое касается наличия телефона мобильного.
  + (42:24) Ну получим так, что по мобильному телефону 100% студентов, 100% студентов и 100% по всей выборке имеют мобильный телефон.
  + (42:34) Теперь по брюкам получим:
  + (42:37) у ребят 100%, у девушек 98%. По всей выборке 99%
  + (42:44) э-э, ходят в брюках.
  + (42:46) Ну у вас там я не вижу, но где-то вот я так немножко прощучиваю, но просто я вам объясняю, как работает модель эта.
  + (42:54) Вот. И длинные волосы:
  + (42:57) э-э, 99% девушек,
  + (43:00) э-э, 1% ребят.
  + (43:03) Вот. В среднем по всей выборке 50 там процентов имеют длинные волосы.
  + (43:08) Значит, смотрим теперь. Э-э, у длинные волосы намного чаще, чем в среднем встречаются в группе девушки, и намного реже, чем в среднем в группе мальчики, да?
  + (43:18) Э-э, а брюки
  + (43:21) чуть-чуть чаще встречаются, чем в среднем у мальчиков, чуть-чуть реже, чем в среднем у девушек.
  + (43:27) Мобильный телефон одинаково, с одинаковой частотой встречается и у мальчиков, и у девочек, и по всей выборке.
  + (43:32) Это означает, что признак длинные волосы очень характерен для девушек. Признак наличие брюк, э-э, характерно для мальчиков, но степень характерности очень небольшая, потому что для девочек тоже он вполне такой часто очень встречается.
  + (43:48) Для девушек этот признак брюки является нехарактерным, но мало содержит информации о том, что это не девушка.
  + (43:55) А вот признак длинные волосы много содержит информации о том, что это не парень, много информации о том, что это девушка. Ну вот эта логика, вот, в общем, я объяснил вам.
  + (44:04) Значит, теперь хочу вам сказать, что если бы матрица модели была побольше, ну, скажем, там 500 колоночек и, допустим, 1.000 строчек, то такое сравнение было бы уже очень трудоёмким.
  + (44:14) То есть уже было бы сложно вот это провести самому сравнение вот этих частот относительных, условных и безусловных.
  + (44:26) Значит, э-э, поэтому я этот процесс автоматизировал. Закрываем эту форму.
  + (44:34) И смотрим... Тем более хотелось бы, э-э, не просто это сравнить, а ещё и решать задачи на основе этого сравнения.
  + (44:45) Вот. То есть надо получить результаты этого сравнения в виде базы данных.
  + (44:48) Как это сделать? Значит, смотрим помощь.
  + (44:51) Кнопочку нажимаем.
  + (44:55) Э-э, на весь экран разворачиваем наш по ширине.
  + (44:59) Видим, ребята, значит, тут я, когда про это рассказываю, я всегда спрашиваю:
  + (45:04) какие способы сравнения двух чисел вы можете предложить? Вот, допустим, у нас есть два числа: 7 и 5. Как узнать, какое из них больше?
  + (45:13) Если вот мы хотим сделать программу, которая будет сравнивать и выводить результат, что 7 больше, чем 5, например.
  + (45:19) Как это сделать можно математически?
  + (45:28) Кто знает?
  + (45:32) Ну, в общем, берём, допустим, эти числа у нас первое и второе. Берём из первого числа вычитаем второе
  + (45:38) и сравниваем. Если разница больше нуля, тогда первое число больше.
  + (45:44) Если наоборот, тогда первое число меньше. А если равно нулю, тогда они одинаковые.
  + (45:51) Согласны?
  + (45:54) Да.
  + (45:55) Хорошо.
  + (45:58) А путём деления можно сравнить эти числа?
  + (46:02) Первое число делим на второе.
  + (46:05) Если результат...
  + (46:07) А?
  + (46:08) Ну да, наверное. Мы же определим часть от числа.
  + (46:13) Конечно. Если больше единицы, тогда числитель больше, чем знаменатель. Если меньше, тогда он меньше.
  + (46:20) А если одинаковое, тогда будет равно единице. Тогда числитель и знаменатель одинаковые.
  + (46:25) Вот. Теперь представьте себе, что вот я вам описал сейчас задачу, когда у нас объект имеет несколько признаков: длинные волосы, брюки, мобильный телефон.
  + (46:33) Мы бы хотели посчитать суммарное количество информации во всех признаках о том, что это, допустим, парень или о том, что это девушка. Нам уже известно, какие эти признаки.
  + (46:43) Понимаете?
  + (46:45) Как мы это сделаем? Мы возьмём и просуммируем это количество информации во всех этих признаках о том, что это парень и о том, что это девушка.
  + (46:53) У нас получится очень большой объём информации о том, что это девушка.
  + (46:58) Понятно, да? Длинные волосы - это очень большой объём информации. То, что брюки немножко его уменьшают, но не намного. А телефон вообще никак не влияет на это.
  + (47:07) Понятно, да? Так вот, для того, чтобы так сделать, именно суммировать, э-э, количество информации в признаках, для этого, это вот называется интегральный критерий, потом я подробнее расскажу про это.
  + (47:19) Для этого нужно, чтобы когда между признаком и принадлежностью объекта классу нет никакой связи, чтобы получалось ноль.
  + (47:27) Вот, допустим, наличие мобильного телефона никакой информации не содержит о том, что это парень или девушка.
  + (47:36) Значит, мы можем вычесть фактическую частоту наблюдения телефона в группе мальчики, в группе девочки и теоретическую частоту.
  + (47:45) И вычесть их друг из друга, и получится у нас ноль.
  + (47:48) То есть фактическая частота и теоретическая равны.
  + (47:52) А можем взять и разделить фактическую частоту на теоретическую, путём деления сравнивать их.
  + (47:58) У нас получится вот выражение, которое мы здесь видим справа.
  + (48:03) Значит, N и Т - это сумма по строке по Итому, N житое - это строка сумма по колонке житой.
  + (48:09) Вот если мы разделим N и житое на N житое, это у нас будет вероятность встречи Итого признака в житой группе, у объектов житой группы.
  + (48:18) Но у нас здесь не вероятности, а частоты. Они стремятся к вероятности при неограниченном увеличении объёма выборки.
  + (48:24) То есть у нас вот это П и Т житое получается. А вот N и Т разделённое на N - это П и Т.
  + (48:30) Это вероятность встретить Итый признак по всей выборке, безусловная вероятность.
  + (48:36) Вот мы их сравниваем путём деления или сравниваем фактическую частоту и теоретическую путём деления. И потом нам нужно, чтобы было равно нулю в случае, когда они одинаковые.
  + (48:47) Это можно сделать двумя способами такими разумными: либо взять логарифм,
  + (48:51) который равен нулю, если у нас под под логарифмом единица находится.
  + (48:56) Либо можно просто взять и вычесть единицу из этого отношения.
  + (49:00) Вот если мы возьмём логарифм, то у нас получается мера количества информации по Харкевичу, ребята.
  + (49:06) То есть это условная вероятность, делённая на безусловную. То есть это логарифм этого отношения. То есть это э-э, количество информации, которое мы получаем,
  + (49:15) э-э, если узнаём, э-э, что у нас вероятность достижения цели повысилась.
  + (49:21) Это по Харкевичу.
  + (49:24) Или понизилась.
  + (49:26) Вот если нам сообщают информацию, что такой-то вот признак у объекта есть,
  + (49:29) и у нас э-э, вероятность... нам стало понятно, что скорее всего это объект относится к группе девушки,
  + (49:36) потому что признак - это длинные волосы.
  + (49:39) То мы получили некоторое количество информации определённое о том, что объект принадлежит к этому классу.
  + (50:03) Об этом Харкевич и говорил, что этот класс, если является целевым состоянием, то это изменяет вероятность достижения целевого состояния, сообщение нам этой информации.
  + (50:12) Ну, примерно как вот, если, допустим, я ищу аудиторию, где надо деньги сдать за обучение.
  + (50:17) И спрашиваю, а где эта аудитория? А мне говорят: вот на первом этаже налево там, э-э, в ФЕ, а там дальше направо.
  + (50:25) Я поворачиваю направо, а там 20 аудиторий там или сколько там их там, 20 там, э-э, скажем, восемь аудиторий, к примеру.
  + (50:33) Тогда, значит, была неопределённость сначала очень высокая,
  + (50:38) одна 256 тысячная, например.
  + (50:41) Вот, или одна там, э-э, 1024, например, одна 1024.
  + (50:46) А когда мне сказали, что вот там в крыле на первом этаже направо, тогда стала 1/8 неопределённость.
  + (50:51) Мы делим 1/8 на 1/1024,
  + (50:55) и получаем, э-э,
  + (50:57) сколько там? 1024 на 8, сколько это будет?
  + (51:06) Вот 128.
  + (51:08) Это 2 в степени 7.
  + (51:12) Э-э, то есть 7 бит информации мы получили,
  + (51:16) когда нам сказали, что вот там эта аудитория.
  + (51:19) А если нам скажут, что это конкретно вот эта аудитория, там 107, к примеру,
  + (51:23) Ну тогда, значит, была 1/1024,
  + (51:29) стала единица. Единица разделить на 1/1024 - это логарифм 1024.
  + (51:34) Это 2 в степени 10. То есть 10 бит информации мы получили.
  + (51:38) То есть чем точнее нам сказали, где эта аудитория, тем больше информации мы получили.
  + (51:44) Это мера Харкевича, ребята.
  + (51:46) Она хорошо работает очень, особенно когда... Во-первых, её легко рассчитать, потому что она основана на таких вот вещах, которые можно из опыта получить, частоты вот эти и вероятности, относительные частоты, условные и безусловные.
  + (52:00) И, во-вторых, она очень удобна для применения в таких системах, где нужно достигать каких-то целей,
  + (52:10) где нужно принять решение о принадлежности объекта классу,
  + (52:13) или нужно принять решение о том, какими факторами мы можем перевести объект в то или иное целевое состояние,
  + (52:19) достичь цели. То есть везде, где есть понятие цели, мера Харкевича очень получается, ну просто вот ей сам Бог велел её использовать.
  + (52:27) Вот так что Александр Харкевич, он предложил замечательную меру.
  + (52:31) Он называл её мера семантическая мера целесообразности информации. Слово целесообразность - это связано с достижением цели.
  + (52:38) Но сейчас современные представления научные говорят о том, что он, по сути дела, предложил меру знаний.
  + (52:43) Потому что информация, которая полезна для достижения цели - это знание уже.
  + (52:47) То есть он был, э-э, одним из первых учёных, может быть, даже вообще первым учёным,
  + (52:52) который предложил количественную меру знаний, понимаете?
  + (52:56) Формальную, не такую вот, когда мы преподаватель там экспертным путём ставит там четвёрку или пятёрку.
  + (53:02) А когда это оценивается по на основе эмпирических данных.
  + (53:08) Строится модель, которая прямо вот содержит определённое количество знаний о предметной области, причём мы можем сказать, какое количество знаний.
  + (53:16) И как мы можем их использовать для решения задач.
  + (53:20) Так что мы, хочу я вам сделать вывод такой сейчас, ребята, что вот эта мера Хи-квадрат, замечательная мера Пирсона,
  + (53:28) она тесно связана с мерой Харкевича количества информации
  + (53:32) и с мерой Рои возврата инвестиций, где просто из этого отношения условно безусловной вероятности вычитается единица, чтобы нормировать их к нулю
  + (53:41) в случае, когда нет зависимости. И в этом случае получается у нас, э-э, что мы можем использовать аддитивные интегральные критерии. То есть можем суммировать вот эти частные критерии,
  + (53:53) и сумму рассматривать как, э-э,
  + (53:57) меру степени соответствия объекта классу.
  + (54:01) Закрываем эту форму.
  + (54:05) И переходим к вопросу об оценке достоверности... Да, смотрим Inf1.
  + (54:10) Модель Inf1. Это прямо количество информации по Харкевичу. Двигаемся вниз,
  + (54:15) вправо.
  + (54:17) Вот как вот мы делали до двадцать третьей этим, до двадцать третьей строчкой.
  + (54:23) Значит, что мы видим, ребята, что в этой двадцать третьей строчке у нас, э-э, мы видим, что, ну да, вправо до конца сдвинулись, да?
  + (54:33) Где тут? Так.
  + (54:36) Двадцать третья строчка.
  + (54:37) Мы получаем 0,779 бита
  + (54:40) о том, что объект относится к средствам связи, зная, что его размер под руку у него.
  + (54:46) И 0,231 бита о том, что это элемент компьютера.
  + (54:51) То есть мы видим, что больше информации о том, что это средство связи, но информация о том, что элемент компьютера тоже есть.
  + (54:58) А вот здесь мы видим с минусом, видите, какие-то числа. Что это значит? Это значит, что эти признаки встречаются у объектов этих классов,
  + (55:05) но реже, чем в среднем.
  + (55:07) Ну это примерно как вот у девушек встречаются, у ребят длинные волосы встречаются, но реже, чем в среднем.
  + (55:14) Признак характерен больше для девушек. Ну бывает у ребят наблюдается.
  + (55:18) Признак характерен для ребят, чтобы брюки. Ну бывает у девушек встречается тоже.
  + (55:23) Тогда это несёт информацию о том, что это не девушка, не парень, соответственно. Понимаете, да?
  + (55:28) То есть смысл в этом минусе именно такой смысл. Закрываем эту форму и смотрим сам Хи-квадрат.
  + (55:35) Я уже сказал, что мера Харкевича и Хи-квадрат взаимосвязаны.
  + (55:39) Смотрим Inf3.
  + (55:43) Это Хи-квадрат.
  + (55:44) Это замечательная сбалансированная модель.
  + (55:47) Кто-нибудь из вас знает, что такое магические, э-э, квадраты?
  + (55:54) Слышали?
  + (55:55) Нет.
  + (55:56) Ну, в общем, это такой квадрат, в котором такие числа в клеточках, что бывают там, э-э, 2х2, бывают 3х3, бывают 100х100.
  + (56:06) Что если их сложить по диагонали, по строке и по колонке, то будет одно и то же число.
  + (56:12) Вот теперь двигайтесь вправо до конца.
  + (56:18) И вниз.
  + (56:20) Сумма равна нулю, среднее равно нулю.
  + (56:23) И вниз тоже сумма равна нулю, среднее равно нулю.
  + (56:27) Это вот свойство меры Хи-квадрат.
  + (56:33) То есть именно по этой причине эта мера очень хорошо работает. То есть вот эти вот плюсы и минусы, они как бы хорошо
  + (56:39) относительно какого-то среднего выравниваются, понимаете?
  + (56:44) То есть их количество такое, что получается, что сумма равна нулю, среднее равно нулю. Это круто вообще-то.
  + (56:50) Вот. То есть получается, что эти модели очень хорошо работают, Хи-квадрат.
  + (56:54) Теперь, э-э, я могу вам сказать, что в магическом квадрате там мог бы быть не ноль, а, допустим, 17, например, там.
  + (57:00) Поняли, да? Во всех клеточках, во всех строчках и по диагоналям.
  + (57:08) Ну почитайте, посмотрите, что такое магические квадраты для прикола, чтобы вы хоть знали про это.
  + (57:13) Закрываем это. То есть я думаю, что модели Хи-квадрат тоже являются частным случаем магического квадрата.
  + (57:20) Только уже здесь не квадрат, а любой формы
  + (57:23) таблица. То есть там может быть разное, не число строк не равно числу колонок.
  + (57:29) Вот, но вот это свойство, оно всё равно сходно вот именно с магическими квадратами, что нулю равно.
  + (57:34) Правда, не какому-то другому значению, а именно нулю.
  + (57:38) Закрываем эту форму, эту модель эту Хи-квадрат.
  + (57:43) Закрываем её. И эту режим 5-5 закрываем.
  + (57:49) Можно Escape нажимать, можно на крестик.
  + (57:52) Теперь давайте переходим, э-э, ну вы уже просто сейчас ещё раз его открываете,
  + (57:58) чтобы запомнить, наверное, да?
  + (58:01) Вот.
  + (58:02) Теперь переходим в режим 3-4.

**4. Оценка достоверности моделей (Ван Рисберген)**

* **4.1. Важность достоверности**
  + (58:08) Так, что мы здесь имеем?
  + (58:16) Режим 3-4.
  + (58:19) Вот. И, значит, что мы здесь видим? Это режим,
  + (58:22) предназначенный для количественного измерения с уровня достоверности модели с помощью различных
  + (58:29) критериев
  + (58:31) достоверности.
  + (58:33) Здесь нет смысла форму эту масштабировать, потому что она с фиксированной
  + (58:37) шапкой левой.
  + (58:41) Вот, поэтому тут, э-э, ну иначе бы было бы непонятно, что там написано.
  + (58:46) Вот, поэтому она не масштабируется. Значит, нажимаем помощь по мерам достоверности, левую кнопочку.
* **4.2. Подход Ван Рисбергена**
  + (59:01) Вот. Значит, э-э,
  + (59:03) есть такой учёный замечательный, Ван Рисберген,
  + (59:06) который, э-э,
  + (59:08) ну, я, честно говоря, в истории не очень силён,
  + (59:12) когда он это предложил, но, по-моему, он предложил это очень давно.
  + (59:15) Он очень преклонном возрасте. Я сейчас, э-э, не сейчас, а некоторое время назад в Википедии смотрел.
  + (59:20) Ещё было написано, что он работает,
  + (59:23) этот учёный.
  + (59:39) Вот.
  + (1:00:17) Ну это его мера, но
  + (1:00:37) Вот. Он, э-э, предложил рассчитывать достоверность модели путём решения задачи идентификации.
  + (1:00:48) Э-э, это очень удачное, разумное решение.
  + (1:00:52) Возникает вопрос: а можно ли оценивать достоверность модели путём решения других задач? Ну, допустим, задачи прогнозирования или задачи принятия решений?
  + (1:01:00) Можно, но удобнее всего с помощью решения задачи идентификации.
  + (1:01:05) Почему? Потому что при идентификации у нас считается, что признаки объекта наблюдаются у него тогда же, в тот же момент времени, когда он относится к определённой категории или классу.
  + (1:01:18) И мы можем, э-э, одну строчку рассматривать в базе какой-то, э-э, во временных, в базе временных рядов, где у нас строки связаны со временем.
  + (1:01:28) Ну, допустим, курсы ценных бумаг или валют.
  + (1:01:32) И вот мы берём какую-то строчечку, у нас вот такие-то показатели в этой строчке там, число банков участвующих в торгах, такие-то там, э-э, такой-то спрос, такое-то предложение, а сбоку у нас прямо курс написан,
  + (1:01:45) э-э, рубля по отношению к доллару, рубля по отношению к евро, доллара по отношению к евро, прямо курсы написаны.
  + (1:01:52) Вот. И спрос там на разные валюты, и предложение разных валют, и число банков участвующих в торгах, и всё это. И мы увязываем это всё. То есть нам для того, чтобы проверить достоверность модели, мы можем взять любую строчку
  + (1:02:05) и попробовать понять по этим вот показателям, которые там справа, какой будет курс, который слева в этой же строчке.
  + (1:02:13) При прогнозировании задача несколько сложнее. То есть мы берём какую-то текущую дату,
  + (1:02:19) ну, берём какую-то строку в этой базе данных, соответствующую какой-то дате, и говорим так: это вот у нас настоящее.
  + (1:02:26) А что было до этого за за неделю до этого, например?
  + (1:02:30) А потом смотрим, что будет через неделю после этого.
  + (1:02:33) И всё это на ретроспективных данных.
  + (1:02:36) И мы применяем модель для того, чтобы оценить, что произойдёт
  + (1:02:40) на следующей неделе от текущей даты по информации о предыдущей неделе.
  + (1:02:46) И у нас получаются некие прогнозы.
  + (1:02:49) Мы проверяем, совпали они или нет, потому что все эти прогнозы по ретроспективным данным, относящимся к прошлому. Всё это уже произошло, и мы можем это сделать. Но сам алгоритм
  + (1:02:58) э-э, решения задачи более сложный.
  + (1:03:02) Одно дело там независимо друг от друга строчки рассматривать, другое дело э-э, смотреть на записи относительно текущей, назад и вперёд,
  + (1:03:12) смотреть там динамику в этих записях и так далее. То есть там неизвестно, что ещё смотреть. Ну, в общем, короче говоря,
  + (1:03:18) вопрос возникает, э-э, гораздо больше вопросов к решению задачи прогнозирования. Сама задача прогнозирования более сложная, чем задача идентификации.
  + (1:03:29) Хотя, в простейшем варианте они очень сходны.
  + (1:03:33) Вот. А задача принятия решений, она ещё сложнее. Почему? Потому что э-э, при идентификации и прогнозировании мы только данные наблюдения используем.
  + (1:03:46) Мы на сам объект моделирования не воздействуем, только на него смотрим, наблюдаем.
  + (1:03:53) Ну если не считать это само само наблюдение воздействием.
  + (1:03:56) Вот. И э-э, делаем на основе этого какие-то выводы.
  + (1:04:00) А при принятии решений мы ещё к тому же и оказываем на него воздействие на этот объект
  + (1:04:06) э-э, управления, в данном случае уже,
  + (1:04:09) объект моделирования и управления.
  + (1:04:11) И э-э, смотрим потом, какие результаты этого воздействия.
  + (1:04:16) Допустим, взяли какую-то дату текущую, взяли предыдущий период,
  + (1:04:21) выработали на основе предыдущего периода рекомендации по управлению объектом моделирования.
  + (1:04:26) Потом взяли и оказали на него соответствующее воздействие.
  + (1:04:32) А потом посмотрели, что получилось.
  + (1:04:35) И вот мы должны иметь информацию об этом, что эти решения были выработаны, эти решения были реализованы,
  + (1:04:42) и мы увидели, какие-то результаты получились, и сравнили их с ожидаемыми результатами, и увидели, насколько достоверная была модель, на основе которой эти рекомендации были выработаны, и увидели, э-э, осуществились они или нет, и поняли, насколько эта модель адекватна, насколько правильно она отражает объект моделирования и его поведение под влиянием наших этих управляющих факторов.
  + (1:05:05) Вот. То есть проверить достоверность модели на основе решения э-э, задачи принятия решений
  + (1:05:11) ещё сложнее, чем на основе задачи прогнозирования.
  + (1:05:15) Потому что нужно воздействие на объект.
  + (1:05:18) Хотя бы в прошлом, но воздействие должно было бы иметь место.
  + (1:05:22) И должны быть результаты известны, которые которые получились при всех этих воздействиях и наших, и окружающей среды.
  + (1:05:33) Поэтому Ван Рисберген совершенно справедливо решил
  + (1:05:36) оценивать достоверность модели по путём решения задачи идентификации.
  + (1:05:41) Что он сделал? Он предложил использовать четыре сумматора... Ну он сперва так: что должна вообще делать модель?
  + (1:05:49) Она должна обеспечивать нам
  + (1:05:52) э-э, правильное отнесение объектов к тем классам, к которым они на самом деле относятся фактически,
  + (1:05:58) и правильное не отнесение объектов к тем классам, к которым они на самом деле фактически не относятся.
  + (1:06:05) И, значит, правильно. Значит, она должна это делать правильно и не должна ошибаться.
  + (1:06:09) Какие виды ошибок бывают в этом случае?
  + (1:06:12) Она должна правильно относить объект к классам, а она может это сделать неправильно.
  + (1:06:17) То есть она может отнести объект к классу, а на самом деле он к нему не относится. Это называется
  + (1:06:23) ложноположительное решение по Ван Рисбергену.
  + (1:06:27) А если система не относит объект к классу, то она тоже должна это делать правильно. То есть если она не относит, то и фактически он не должен к нему относиться к этому классу объект.
  + (1:06:38) А на самом деле он может к нему относиться.
  + (1:06:42) Тогда система, значит, ошиблась, приняв не отнесение его к классу.
  + (1:06:48) Это ложноотрицательное решение.
  + (1:06:50) Значит, он ввёл термины соответствующие. Ну, ложно истинно - это понятно, True False.
  + (1:06:55) А вот решение о принадлежности объекта классу - это положительное решение.
  + (1:07:00) Решение о непринадлежности объекта классу - это отрицательное решение.
  + (1:07:05) Как это определяется? Положительное решение или отрицательное? По знаку уровня сходства объекта с классом.
  + (1:07:11) Если знак уровня сходства положительный, тогда это называется положительное решение. То есть объект похож на класс.
  + (1:07:18) А если объект не похож на класс, тогда знак с уровня сходства отрицательный.
  + (1:07:22) И получается, что это решение отрицательное.
  + (1:07:28) То есть негативное, позитив, негатив.
  + (1:07:30) Позитив - положительное, негатив - отрицательное.
  + (1:07:33) Соответственно, Ван Рисберген ввёл четыре сумматора:
  + (1:07:36) True Positive, True Negative, False Positive, False Negative.
  + (1:07:42) И предложил посчитать значение этих сумматоров при решении задачи идентификации.
  + (1:07:48) И потом, э-э, предложил формулы, обосновал их,
  + (1:07:52) по которым он рассчитывает точность модели, полноту модели и достоверность модели.
  + (1:07:58) При этом очень много разных вариантов
  + (1:08:00) F-меры Ван Рисбергена, F-бета-меры,
  + (1:08:03) для случая, когда у нас только два исхода есть, когда много вариантов исхода.
  + (1:08:11) И есть симметричные, несимметричные меры, когда больше играет роль точность, когда больше играет роль полнота и так далее, и так далее. Я видел,
  + (1:08:19) ребята, статьи или научные работы, где целая страница исписана разными формулами,
  + (1:08:25) связанными с критериями Ван Рисбергена и способами
  + (1:08:30) там оценки различных параметров модели.
  + (1:08:33) То есть их довольно большое количество этих параметров.
  + (1:08:36) Но в простейшем варианте это всё легко найти в интернете, поэтому я по этому поводу не распространяюсь особенно.
  + (1:08:42) Могу единственное сказать, что вот их очень много таких, э-э, моментов, деталей.
  + (1:08:49) Но суть такая, что очень простая она. Рассчитывается на основе этих сумматоров, э-э, точность и полнота модели,
  + (1:08:58) и потом как среднее гармоническое рассчитывается её достоверность модели.
  + (1:09:03) И вот смотрите, ребята, значит, вот здесь у нас есть эти формулы приведены.
  + (1:09:09) Классический критерий Ван Рисбергена, общепризнанный, убедительный.
  + (1:09:15) Значит, я вам очень советую, ребята, вы же, э-э, экономисты, да, я так понимаю?
  + (1:09:21) Вы, э-э, будете связаны с моделированием, потому что всё всё равно, когда мы проводим научные исследования, то мы строим модели объекта исследования,
  + (1:09:30) познаём этот объект исследования, отражаем результаты его познания в модели некоторой.
  + (1:09:37) Или сама модель является результатом процесса познания, или может даже инструментом познания.
  + (1:09:43) И вот, э-э, вы должны хорошенько запомнить одну вещь. Я вот просто вас призываю это запомнить.
  + (1:09:49) Как только вы слышите слово "модель",
  + (1:09:51) вы сразу же должны вспоминать про слово "достоверность".
  + (1:10:00) Кто-то говорит: "Вот, вот там модель, модель..." Вы только раз, он сказал "модель". А Евгений Венеаминович говорил, что как только он скажет "модель", сразу должна следующая фраза там или скоро должна услышать что-то про достоверность этой модели,
  + (1:10:12) как он её оценивал.
  + (1:10:13) И вы слушаете, слушаете, а он понесло, понесло его дальше, дальше. И он говорит: "Вот мы эту модель применяли для того, чтобы выработать рекомендации, там выработали рекомендации, там и так далее, и так далее".
  + (1:10:24) А слово "достоверность" не прозвучало вообще.
  + (1:10:27) И у вас возникает такой вопрос: а он какое право имел вырабатывать рекомендации,
  + (1:10:32) не зная, э-э, достоверна модель или нет?
  + (1:10:38) Вы представьте себе, что модель его могла оказаться недостоверной.
  + (1:10:43) Она ж могла быть не только достоверной, но и недостоверной.
  + (1:10:45) Тогда бы он выработал бредовые рекомендации. Эти бредовые рекомендации применили бы на практике, и фирма попала бы в банкротство, допустим, или тяжёлое состояние.
  + (1:10:54) Огромные потери были бы финансовые, а может быть, даже и не только финансовые.
  + (1:11:00) А ведь разные ж решения принимаются на основе модели, и в боевых условиях принимаются решения.
  + (1:11:06) Тогда ошибочные решения выражаются в тысячах погибших людей,
  + (1:11:10) десятках тысяч травмированных.
  + (1:11:14) Возникает вопрос: а нас какое вообще ответственность человек несёт за принятие неправильных решений?
  + (1:11:21) За то, что он, э-э, применял какие-то неразумные модели,
  + (1:11:25) которые неверно отражают реальность,
  + (1:11:28) и он на основе этих моделей принимал решения, которые повлекли, э-э, колоссальные отрицательные последствия.
  + (1:11:35) Он вообще какое моральное право имел применить эту модель? Я вам могу сказать: никакого, никакого, нет никакого оправдания в этом.
  + (1:11:42) То есть он должен был подумать, насколько его решения адекватны, насколько его модели правильно отражают реальность.
  + (1:11:49) Насколько ожидаемый результат, э-э, желаемый, желаемый им, э-э, реально может быть достигнут таким путём.

**5. Заключение и дальнейшие шаги**  
\* (1:11:59) Если модель, ребята, не измерена на достоверность,  
\* (1:12:09) и неизвестно, какая у неё достоверность,  
\* (1:12:12) то нельзя решать никакие задачи с помощью этой модели: ни задачи идентификации, ни задачи прогнозирования,  
\* (1:12:18) ни задачи принятия решений,  
\* (1:12:20) ни задачи исследования моделируемой предметной области.  
\* (1:12:23) Вот если эта модель неправильно отражает объект моделирования, мы берём на решаем задачу исследования моделируемой предметной области путём исследования модели.  
\* (1:12:31) Получаем некие результаты. Эти результаты никакого отношения к реальности не имеют, потому что модель не отражает реальность, понимаете?  
\* (1:12:39) Вот. То есть исследовать бесполезно модель объект моделирования путём исследования модели, которая не отражает объект моделирования.  
\* (1:12:47) Ребят, сколько вас там человек осталось?  
\* (1:12:52) Или не ребята, а девушки?  
\* (1:12:53) Нас так же четверо.  
\* (1:12:54) Четыре. Понятно.  
\* (1:12:58) Просто двое отключились.  
\* (1:13:03) Такое впечатление, значит, ну ладно.  
\* (1:13:06) То же самое касается идентификации.  
\* (1:13:07) Простите, извините, интернет плохо работает, и поэтому закрыли некоторые. Ну, меня выкинуло просто с телефона.  
\* (1:13:16) Ясно. Ну вы вообще-то меня хоть слышите, что я рассказываю, нет?  
\* (1:13:20) Да. Смотрите на эту экранную форму, да? Help? Ага.  
\* (1:13:25) Так вот, смотрите, ребята, значит, э-э, Ван Рисберген, ну я так предполагаю, что он, наверное, программист, потому что он ввёл вот эти вот такие обозначения для этих сумматоров: ТП, ТН, ФП, ФН.  
\* (1:13:38) Потому что математики, у них принято одной буквой обозначать какую-то величину.  
\* (1:13:46) Ну ещё и индексы они используют.  
\* (1:13:48) А вот такие двубуквенные обозначения, они бывают у экономистов и у программистов. Для программиста это совершенно естественно выглядит вот такое введение таких обозначений.  
\* (1:14:02) Вот. И вот он ввёл эти обозначения, предложил их посчитать,  
\* (1:14:06) и на основе них предложил эту меру свою.  
\* (1:14:10) Так.  
\* (1:14:11) Так что мы, хочу я вам сделать вывод такой сейчас, ребята, что вот эта...  
\* (1:14:22) Так, ребят, значит, давайте на этом, значит, мы занятие наше заканчиваем.  
\* (1:14:25) Э-э, я у себя отмечу в расписании, что на следующем занятии мы продолжаем рассмотрение вопроса достоверности модели.  
\* (1:14:36) И на этом у нас занятие заканчивается.  
\* (1:14:44) До свидания.  
\* (1:14:46) Спасибо. Можно выходить, да?  
\* (1:14:49) Да, да, да, да, конечно.  
\* (1:14:52) До свидания. Всего самого...  
\* (1:14:54) До свидания.  
\* (1:14:55) Счастливо.  
\* (1:14:56) Счастливо.