***ФГБОУ ВО «Кубанский государственный аграрный университет имени***

***И.Т. Трубилина»,***

***Российская Федерация***

***Колесников Роман Юрьевич, ПИ2102***

***roman563412@gmail.com***

**Инсталляции. Облачное Эйдос-приложение.**

**Заголовок**

Система Эйдос: Типы приложений, лабораторных работ и обработка данных

**Резюме**

В видеоролике представлен обзор программной системы Эйдос, предназначенной для выполнения лабораторных работ и анализа данных. Рассматриваются два основных типа приложений: локальные (устанавливаемые) и облачные (доступные через FTP-сервер).

1. Типы приложений и лабораторных работ:

Локальные приложения: Устанавливаются на компьютер пользователя, доступны через инсталляцию. Пример добавления – режим 1-3, кнопка "Добавить лабораторную работу".

Облачные приложения: Находятся на FTP-сервере, скачиваются из облака (режим 1-3, кнопка правее). Существует каталог из 320+ облачных приложений, многие с описаниями. Облачные приложения образуют форум, где каждую работу можно обсуждать.

Типы лабораторных работ:

Тип 1: Готовые к установке и изучению работы. Теория доступна онлайн с сайта автора.

Тип 2: Исходные данные формируются расчетным путем внутри системы, не требуют внешних данных или интернета (кроме скачивания описания).

Тип 3: Используют внешние исходные данные (текстовые, числовые) из файлов, например, Excel. Демонстрируют обработку таких данных.

Тип 4 (и дальнейшие режимы 2-3-2-X): Ориентированы на интеллектуальную обработку текстов и изображений (графических данных) различными способами (по контурам, пикселям, спектру).

2. Работа с данными и интерфейсами:

Ввод данных: Система поддерживает ввод данных из различных источников, включая таблицы Excel. Демонстрируется интерфейс 2-3-2-2 для ввода табличных данных.

Структура данных: Система оперирует "наблюдениями" (объекты, состояния, процессы). Каждое наблюдение описывается двумя типами шкал:

Описательные шкалы: Характеризуют свойства наблюдения (вес, цвет) или влияющие факторы (температура, давление). Могут быть числовыми, номинальными (текстовыми, неупорядоченными), порядковыми (текстовыми, упорядоченными).

Классификационные шкалы: Относят наблюдение к определенным категориям или классам (будущим состояниям).

Обработка данных: Система преобразует все разнородные данные (числовые, текстовые, разных единиц измерения) в единую меру – количество информации, обеспечивая их сопоставимость. Это достигается через создание системно-когнитивных моделей.

Интерфейс ввода (Пример 2-3-2-2): Позволяет настраивать параметры считывания данных из Excel (старый/новый формат, DBF, CSV), обработку нулей и пробелов, создание баз данных средних значений, указание диапазонов классификационных и описательных шкал (колонок в Excel). Числовые шкалы могут разбиваться на интервалы (равные или с равным числом наблюдений).

3. Синтез и верификация моделей:

Процесс: После ввода и формализации данных (создания справочников шкал, кодирования данных, создания обучающей выборки) система переходит к синтезу и верификации моделей.

Модели: Создаются статистические (матрица сопряженности, условных/безусловных частот) и системно-когнитивные модели (на основе информации, ХИ-квадрат и др.). Модель ХИ-квадрат обладает свойствами, схожими с магическим квадратом (нулевые суммы и средние по строкам/колонкам).

Верификация: Оценка достоверности моделей критически важна перед их использованием для решения задач (идентификации, прогнозирования, принятия решений). Верификация выполняется путем решения задачи идентификации на обучающей выборке (или ее части – бутстреп-подход). Нельзя использовать непроверенные модели для выработки реальных рекомендаций из-за риска неадекватности. Достоверность измеряется различными критериями (F-мера, L1, L2), нормированными от 0 до 1.

4. Практический пример:

Демонстрируется добавление локальной лабораторной работы 3.03 (третий тип).

Показывается процесс настройки параметров ввода данных из Excel через интерфейс 2-3-2-2.

Запускается процесс синтеза и верификации моделей, отображаемый прогресс-баром.

Объясняется структура файлов (папка LabWorks с данными для разных типов работ, папка InpData для текущей работы).

Детальная расшифровка текста

1. Введение в типы приложений

1.1 Локальные и Облачные приложения

Ну там все эти приложения были в инсталляции.

И второй тип приложений - это облачное Эйдос-приложение, которые находятся на FTP сервере.

Вот. Значит, вы можете скачать локальные приложения, которые в инсталляции, нажав кнопочку "Добавить лабораторную работу" в режиме 1-3.

Вот, а облачное сейчас я вам покажу.

2. Типы Лабораторных работ и работа с ними

2.1 Обзор типов лабораторных работ

И здесь у нас есть три, то есть четыре даже типа лабораторных работ.

Тип 1:

Работы первого типа - это работы, которые можно установить и изучать. То есть они полностью готовы для освоения, для изучения.

Тип 2:

Работы второго типа - это работы, э-э, у которых э-э, исходные данные формируются расчетным путем.

То есть ничего не требуется извне. Система сама посчитает исходные данные, и можно будет эти лабораторные работы изучать.

То есть независимо от того, есть интернет, нет интернета.

Тип 3 и 4 (Обработка внешних данных, текстов, изображений):

(Переход к типу 3) Берём работу третьего типа, смотрим. И здесь у нас есть... меняется название работы, естественно, вот в этом окошке, какие работы относятся к этой категории. И здесь у нас есть работа 3.03. Э-э, это работа демонстрирует и показывает, как использовать исходные данные внешние, находящиеся в экселевском файле. Это могут быть и текстовые данные, и числовые.

И также мы можем изучать, и мы будем это делать, изучать, как обрабатываются тексты интеллектуально в системе Эйдос, и как обрабатываются э-э, графические э-э, изображения. Э-э, тоже интеллектуальная обработка изображений.

Это всё есть у нас э-э, в режиме 2 3 2 2. Интерфейс - это табличный интерфейс. А 2 3 2 1 - это текстовые э-э файлы считываются для обработки, интеллектуальной обработки текстов.

А вот дальнейшие 2 3 2 3, 2 3 2 4 - это всё обработка интеллектуальная обработка изображений разными способами. Обработка по контурам, по всем пикселям и спектру. Вот. Э-э, то есть это разные интерфейсы ввода изображений. Мы тоже это потом с вами изучим. Планируется это сделать.

2.2 Доступ к теории и описаниям

Хотя вообще, э-э, описание лабораторных работ находятся на моем сайте, и для того, чтобы их почитать, нужен интернет.

Вот, допустим, берём, ставим курсор на работы первого типа, нажимаем кнопочку "Теория по лабораторной работе номер 101", клацаем по этой кнопочке, и с моего сайта загружается описание этой лабораторной работы.

Это просто раздел учебного пособия 2006 года. Вот видно из строки адреса, что это так.

3. Облачные приложения и Форум Эйдос

3.1 Скачивание и обзор облачных приложений

Вот. Теперь, что касается облачных приложений. Тоже в режиме 1-3 мы можем нажать кнопочку, которая правее находится, чем добавить лабораторную работу. Правее есть кнопочка "Скачать приложение из облака".

И здесь мы видим, э-э, скачивается каталог этих приложений. На данный момент, э-э, есть 320 облачных Эйдос-приложений с описаниями, почти все с описаниями. Ну не все на самом деле, не все, если честно. Мы можем это посмотреть, где там есть описание. Вот синенькая колоночка на синем фоне, видите, там почти всё полностью заполнено. Но некоторые отдельные не заполнены, вот там описания нет, значит, этих работ.

Любую лабораторную работу из облака, или, вернее, любое приложение из облака, мы можем почитать, кликнув два раза на этом, на этой строчке. Тогда сразу же скачивается описание этого приложения, которое находится тоже в облаке.

Мы видим описание.

3.2 Установка облачных приложений

И мы можем это приложение установить, поставить курсор на него и установить, э-э, нажав кнопочку "Установка Эйдос приложения". И тогда это приложение будет установлено. Будет скачано из облака его описание, исходные данные, определён тип программного интерфейса, который нужно использовать для ввода этих данных, API автоматизированного программного интерфейса. Он будет запущен с нужными параметрами, они тоже записаны в облаке. И это приложение будет установлено.

3.3 Обсуждение облачных приложений (Форум)

Значит, да, ещё вот хотел сказать про облачное приложение, я сейчас это упустил, что каждое из них является темой форума. И и каждое из них можно обсуждать, поставив курсор на это приложение и нажав кнопочку справа "Обсуждение этого приложения".

Можно его обсуждать с кем угодно, кто будет делать то же самое, или с автором этого приложения, или со мной, или с другими э-э, посетителями, пользователями системы Эйдос. То есть это вот список этих облачных приложений, по сути дела, это форум, в котором каждое приложение является темой форума.

4. Практический Пример: Добавление и Конфигурация Лаб. Работы 3.03

4.1 Добавление локальной работы

Но сейчас мы будем изучать э-э, локальное учебное приложение 3.03. Ставим курсор на кнопочку "Добавить лабораторную работу", выбираем лабораторную работу третьего типа и клацаем.

Ну а теперь давайте в режиме 1-3 выбираем третий тип приложения. Добавить лабораторную работу, третий тип, третью работу 3.03 и нажимаем Окей.

Сразу появляется сообщение, что сейчас все файлы в папочке, где находятся системы, и Data, и InpData будут удалены.

4.2 Структура файлов и папок для данных

Что это за папочка, ребята? Значит, это папочка, которая предназначена для исходных данных, для ввода исходных данных в систему Эйдос.

А откуда же будет э-э, установлены исходные данные? Значит, у нас, вы видели, 30 лабораторных работ, по 10 лабораторных работ каждого типа. Вот есть вот папочка такая LabWorks. В папочке АИдаТАС есть папочка LabWorks, что можно условно перевести как лабораторные работы.

Значит, вот возьмём сейчас первые 10 лабораторных работ - это работы первого типа. С одиннадцатой по двадцатую работу - работы второго типа, с двадцать первой по тридцатую - третьего типа.

И вот мы сейчас рассматриваем третью работу третьего типа, то есть работу 23. Вот если мы посмотрим сейчас на файл name LW LabWorks, то мы увидим название этой лабораторной работы. И здесь же мы видим файл InpData, видите? И файлик 2 3 2 2 Arch. Это файл с параметрами программного интерфейса ввода этой лабораторной работы.

4.3 Загрузка данных и запуск интерфейса

Если сейчас я нажимаю здесь вот Окей, то все файлы в этой папочке исходных данных стираются, и туда записывается файл из соответствующей папочки э-э, соответствующей лабораторной работы.

И открывается окно э-э, автоматизированного программного интерфейса 2 3 2 2. Это универсальный программный интерфейс ввода данных из таблиц, из табличных данных, из Экселя.

4.4 Проверка интерфейса пользователями

Значит, у вас сейчас должно появиться вот это окошко этого программного интерфейса на экране. Пожалуйста, у кого оно появилось, значит, ну, нажмите, ну поднимите руку, чтоб я это увидел, что у вас это получилось. У кого? Тут у вас такая толпа, что не видно прямо, кто руку поднял. Значит, уже двое подняли: Екатерина, Виктория. А Анастасия, Ангелина почему не поднимают? Анастасия тоже подняла. Ангелина с телефона, да, я так что-то?

Да, я с телефона.

Ну тогда смотри просто, как это делается. И, в принципе, ты можешь потом попытаться то же самое сделать. Запись идёт, я так смотрю. То есть ты, в общем, можешь это же самое всё сделать. Или без записи тоже, в принципе, можешь сделать.

4.5 Обзор интерфейса ввода данных (2-3-2-2)

Значит, здесь довольно много разных параметров. Все эти параметры по умолчанию заданы. То есть, э-э, вместе с исходными данными скачан файл параметров. Можете опустить руки. Значит, и они здесь вот у нас и показаны в этой экранной форме эти параметры. То есть менять их не надо. Но я про них расскажу про эти параметры.

Значит, мы здесь видим слева э-э окошко, где тип данных задаётся внешне. Значит, здесь вот у нас есть старый Excel, новый Excel, DBF, CSV файлы.

Справа вверху окошки у нас написано, где мы задаём, считать ли нули и пробелы отсутствием данных или считать значениями данных. И здесь есть птичка создавать ли базы данных средних по классам. Кому-то это может быть интересно, но обычно это не используется, но можно создавать.

И потом указывается слева диапазон классификационных шкал, которые мы будем рассматривать в экселевской таблице шкалы как классификационные. И диапазон описательных шкал. Какие, какой диапазон шкал мы будем рассматривать как описательные. Шкалами являются колонки исходного экселевского файла, файла исходных данных.

Ниже задаётся ещё ряд параметров. Сейчас я про них расскажу вам. А сначала довольно долго буду рассказывать вам о том, э-э, какой стандарт данных экселевских в этом интерфейсе 2 3 2 2.

И тут вы видите, Help есть вверху на самом этом, на самой экранной форме с параметрами. Есть вверху Help. И этот Help э-э содержит подробнейшее описание всего, что нужно. И у него есть две странички у этого хелпа. Внизу первого хелпа есть кнопочка, нажав которую, можно вызвать вторую страничку. На второй страничке в основном содержатся определения, но не только, там и некоторые смысловые вещи.

5. Стандарты и структура входных данных

5.1 Понятие "Наблюдение"

Вот сейчас я вам расскажу, как устроен этот файл исходных данных. Ну, прежде всего, вы должны знать, что в системе Эйдос обрабатываются наблюдения. Наблюдениями могут быть либо объекты, либо состояния объектов или процессов и явлений, или процессы и явления. То есть, вообще говоря, понятие наблюдения очень широкое. Это может быть практически всё, что угодно. Это могут быть состояния людей, состояния предприятий, состояния полей, вот, состояния там каких-то химических реакций, процессов. Вот, э-э, состояние спутников, вот, или каких-то машин, э-э, которые мы можем определить, э-э, должны определить, что у неё за состояние по параметрам телеметрии. То есть, может быть, самые разнообразные какие-то, может быть, э-э, что угодно практически в качестве наблюдения.

5.2 Описательные и Классификационные шкалы

Вот, но каждое наблюдение описывается двумя способами. Вот здесь вот уже существенный момент. С одной стороны, описывается наблюдение э-э, значениями э-э, градациями описательных шкал, а с другой стороны - градациями классификационных шкал. Вот это то, что сейчас сказал, это очень общее понятие, очень общее определение, но оно, видите, не очень понятное, потому что непонятно, что такое описательные шкалы и классификационные шкалы.

Описательные шкалы позволяют нам описывать наблюдения, а классификационные шкалы позволяют относить это наблюдение к тем или иным категориям обобщающим.

5.3 Свойства и Факторы (в описательных шкалах)

И как же мы можем его описывать это наблюдение? Мы можем его описывать либо значениями свойств, либо э-э значениями факторов, которые влияют на объект наблюдения, объект моделирования. То есть колоночка соответствует либо свойству, либо фактору.

Свойство имеет значение. Ну, например, свойство вес имеет значение 60 кг в одном наблюдении и 70 кг в другом наблюдении. А значение цвет в одном наблюдении имеет значение красное, а в другом значение наблюдение имеет значение зелёный. Понятно, да, ребята? То есть свойство и их значение.

То же самое касается факторов. Факторами являются свойства обычно технологии или окружающей среды, влияющих на объект моделирования. Ну, например, температура, давление. Вот. Вот, допустим, температура 40° там по Цельсию. Это фактор, который определённым образом влияет на объект моделирования. Вот такой фактор влияет с таким значением, и мы видим, что получается с объектом моделирования. И это то, что получается, отражено в классификационных шкалах.

5.4 Статическая и Динамическая интерпретация данных

Если у нас э-э описательные шкалы - это свойства и их значения, то классификационные шкалы, градации классификационных шкал - это обобщающие категории, к которым относятся состояния объекта моделирования.

Если же у нас э-э описательные шкалы - это факторы, то обычно факторы действуют в прошлом периоде по отношению э-э к э-э категориям классификационных шкал, то есть градациям классификационных шкал, классам. То есть классы отражают будущее состояние объекта моделирования или управления, в которое он переходит под действием вот этих значений факторов.

Понятно, да? То есть у нас есть статичная интерпретация данных, когда э-э мы как описательные шкалы и градации рассматриваем свойства и их значения, а как классификационные рассматриваем э-э текущие состояния объекта моделирования, наблюдаемые в то же время, когда у них наблюдаются те вот или иные свойства и их значения.

И динамическая интерпретация, когда описательные шкалы градации описывают факторы, действующие в прошлом, по отношению к э-э градациям классификационных шкал, классам, которые описывают будущее состояние объекта моделирования управления, в которое он переходит под действием этих вот факторов.

Есть универсальная классификация, в которой не конкретизируется, э-э как они относятся во времени, э-э градации описательных шкал и классификационных. И просто используется такая терминология, что градации описательных шкал в какой-то степени характерны, не характерны для объектов, которые относятся к тем или иным классам.

5.5 Типы шкал (Числовые, Номинальные, Порядковые)

Значит, шкалы и описательные, и классификационные могут быть и различных типов данных. Значит, э-э это данные числовые могут быть в этих шкалах и текстовые. Если шкалы текстовые, то они могут быть либо номинальные, либо порядковые. В номинальных шкалах отношение только эквивалентности, не эквивалентности есть. То есть у объекта либо есть определённое свойство, либо нет, либо он относится к какому-то классу, либо не относится. В порядковых шкалах, значит, есть э-э степень выраженности свойства. Вот. И, значит, э-э уже отношение есть не только эквивалентности, не эквивалентности, но ещё и отношение больше-меньше.

И наиболее высокую степень формализации обеспечивают числовые шкалы, в которых, кроме всего предыдущего, отношения эквивалентности, не эквивалентности, больше-меньше, ещё существуют отношения, которые можно выразить арифметическими операциями: сложение, вычитание, деление и умножение. Те из вас, кто знакомы с понятиями математики, кольцо, поле, они, наверное, вспоминают, что что-то похожее слышали при определении этих э-э понятий математических множеств таких типов. Это как раз определение множеств, над которыми возможны такие операции, которые я сейчас перечислял.

И, соответственно, степень формализации модели будет тем выше, чем больше, чем более высокоформализованные шкалы используются. Значит, номинальные шкалы позволяют обрабатывать текстовую информацию. Э-э, порядковые тоже текстовую, но уже с отношениями порядка. А числовые, естественно, понятно, информацию в различных единицах измерения.

5.6 Сопоставимость данных и Нормализация в информацию

То есть мы можем в системе Эйдос использовать шкалы различных типов и в различных единицах измерения. И она обеспечивает решение задачи сопоставимости обработки наблюдений в различных типах шкал и различных единицах измерения.

Эта проблема сопоставимости решается тем, что все э-э наблюдения преобразуются на основе всех наблюдений создаются матрицы системно-когнитивных моделей, в которых э-э все значения всех факторов или все значения всех свойств и значения всех классов э-э рассматриваются лишь с одной точки зрения: какое количество информации содержится в том или ином значении фактора или в том или ином значении свойства о том, что объект принадлежит или перейдёт в определённое состояние, соответствующее классу.

Раз всё преобразуется в информацию, то всё обрабатывается сопоставимо в одной единице измерения. Мы же не можем складывать сантиметры с километрами или сантиметры с килограммами тем более. Нужно перевести в одну единицу измерения. Ну как перевести сантиметры и дециметры в одну единицу ещё понятно или там метры, километры. А вот как перевести в одну единицу килограммы и сантиметры не очень понятно. В системе Эйдос это тоже делается. Это делается путём перевода э-э этих всех э-э значений в количество информации, которое в них содержится о переходе объекта моделирования в те или иные состояния. Это очень важная особенность системы Эйдос и её одно одно из её достижений, изюминок, я бы сказал бы я.

5.7 Настройка интервалов числовых шкал

Значит, дальше мы для числовых шкал указываем способ выбора размера числовых интервалов. Э-э, значит, э-э, либо это у нас равные интервалы с разным числом наблюдений, либо разные интервалы с равным числом наблюдений. Но равным э-э до целых чисел. Ну, примерно, как э-э у нас задача о землекопах известная, знаменитая. Если два землекопа копают две ямы, то сколько землекопов копают одну яму? Ну, берём 3 делим на 2, получается полтора. И э-э третьеклассники там или пятиклассники говорят: "Полтора землекопа". Даже мультик, по-моему, есть про это. Вот. Ну все нормальные люди понимают, что полтора землекопа не будет копать никаких ям. Вот, а один будет копать, а другой, я не знаю, что он будет делать, если его половина там, например. Так вот, э-э, получается, ясно совершенно, что яму копает э-э либо одну яму копает один землекоп, а другую два. Ну, в среднем полтора, да, действительно. Но это это в среднем, оно какой-то смысл имеет целочисленный в данном случае. Так же точно и наблюдения у нас целое количество наблюдений. У нас не может быть там полтора наблюдения. Поэтому у нас, если интервалы разных размеров, то тоже э-э на один на одно числовое, на один числовой диапазон, интервал приходится какое-то целое число наблюдений. Ну, допустим, один или два. Ну, то есть целое, не полтора.

5.8 Дополнительные параметры интерфейса (Сценарии, Обработка текста)

И здесь у нас в этом программном интерфейсе ещё есть много разных возможностей. Можно применить сценарный системно-когнитивный анализ, где учитывается не только точечное значение, но и динамика их изменений в прошлом и будущем, на заданную э-э глубину предыстории и горизонт прогнозирования.

Можно э-э рассматривать поля текстовые как содержащую э-э текст, который будет целиком анализироваться, или только слова будут анализироваться, или только элементы слов. И при этом у нас будут либо, если слова анализируются, либо будет использоваться лемматизация, либо не будет использоваться. Мы это будем рассматривать э-э на занятиях, когда будет у нас лабораторная работа по интеллектуальной обработке текстов.

Когда будут формироваться шкалы и градации, то мы можем задать один из вариантов э-э наименований градаций. Либо числа числовой чисто числовой диапазон. Но если текстовая шкала, то там градациями являются просто уникальные текстовые значения. Если же это градация, то есть если шкала числовая, то градациями являются числовые диапазоны. И можно просто их указывать в качестве наименования. Ну, допустим, один наклонная черта три, и дальше минимальное, максимальное значение этого числового диапазона. А вот эти числа там один наклонная черта три, они означают, что всего у нас три числовых диапазона в этой шкале, а это первый диапазон, минимальный. А можно добавить, чтобы было слово написано перед числовым диапазоном: минимальный, средний, максимальный. И там и разные варианты есть: минимальный, э-э маленький, средний, большой и максимальный, например, или очень большой. Итак, есть тогда большого числа диапазонов разные слова используются, которые имеют соответствующий смысл. Либо можно только слова использовать, а числовой диапазон не указывать. Ну я считаю, что наилучшим вариантом является первый, потому что там наиболее короткие наименования получаются, и смысл их понятен вполне за счёт того, что здесь указано число диапазонов и какой диапазон конкретно имеется в виду.

6. Синтез и Верификация Моделей

6.1 Процесс формализации и создания моделей

Значит, теперь мы этот программный интерфейс в чат посылаем, чтобы у вас он был, хотя я не помню, может, там он уже есть. И, значит, э-э экранная форма программного управления программным интерфейсом. Э-э, и мы нажимаем теперь Окей. И у нас происходит ввод файла исходных данных в систему и анализ его. И результат анализа выводится вот такой экранной форме, которую я называю внутренний калькулятор.

6.2 Внутренний калькулятор (Результаты анализа)

Значит, сейчас, ребята, пожалуйста, э-э отметьте, э-э поднимите руку, у кого получилось вот то, что я сейчас показал. Кто вышел на эту вот экранную форму внутреннего калькулятора? Вот у Виктории прямо в реальном времени всё получается. Как я показываю, так у тебя и получается, да? Поэтому ты такая довольная на фотке. А ещё у остальных как? Ну, Анастасия, я понял, что телефон, да? Или Ангелина, у кого там? Вот. А Екатерина, Анастасия, как? Получилось у вас выйти на этот режим калькулятора? Или возникли проблемы какие-то? У Екатерины тоже получилось, да? У Анастасии получилось. Хорошо. Значит, проблема бывает, знаете какая? У некоторых на компьютерах не установлены офисы. Вот. И тогда, если там, допустим, Open какой-нибудь офис или что-нибудь типа того, значит, э-э то Excel то есть там конвертер экселевских файлов в DBS формат отсутствует, потому что он есть именно в офисе. И тогда э-э система выдаёт соответствующее сообщение об этом. Не об ошибке, а именно сообщение, что у вас там установлена версия, в которой нет этого конвертера. И советуют, что установить. Значит, иногда бывает установлена версия Экселя новая какая-нибудь, там шестнадцатый, девятнадцатый год, в которой тоже этого конвертера нет. Иногда бывает, что он есть, кстати, в новых версиях Экселя. В некоторых версиях он почему-то есть, а в некоторых нет. Не знаю, профессиональная, непрофессиональная. Ну, в общем, короче, это дело вылавливается и сообщение выдаётся.

Значит, здесь мы видим результаты анализа, что у нас есть одна числовая описательная шкала, видите? Числовая одна. Справа описательные шкалы, слева классификационные. Восемь текстовых, а здесь две текстовых в классификационных шкалах.

И вот, поскольку здесь есть в описательных шкалах числовая шкала, то задаётся запрос: а сколько интервальных значений разбивать эту шкалу? И мы можем задать, допустим, на три интервальных значения. Тогда надо пересчитать. Обратите внимание, здесь мы видим 14 колонок, соответствующих классам, градациям классификационных шкал. И 49 э-э строчек, соответствующих градациям описательных шкал.

Вот я сейчас нажимаю пересчитать, у нас получается, смотрите, 40 уже, видите, э-э строчек, соответствующих градациям описательных шкал. Ну я восстановлю то, что было.

6.3 Переход к синтезу и верификации

Вот. И мы переходим дальше э-э на создание модели.

Переходим на создание модели. У нас появляется экранная форма, где прогресс-бар есть, и э-э прогресс-бар показывает стадию исполнения, а, значит, и прогноз исполнения, когда закончится этот процесс. На нашей этой маленькой модели этот процесс закончился мгновенно.

Вот. Ну он может продолжаться несколько минут на больших моделях, там, ну может минут 5, 7, 10 минут, может, где-то такие вот интервалы времени.

6.4 Выполненные работы при формализации

Что произошло, ребята? Значит, произошло, какие выполнены работы? Первое. Сформированы классификационные и описательные шкалы градации на основе файла исходных данных. То есть был проанализирован файл исходных данных и созданы справочники классов, то есть будущих состояний объекта объекта моделирования, и справочники описательных шкал градаций, то есть справочники факторов и их значений или свойств и их значений.

Второе, что было сделано. Э-э, эти справочники были использованы, то есть классификационно-описательные шкалы градации были использованы для кодирования файла исходных данных. В результате этого кодирования была создана обучающая выборка или я её называю ещё базой событий. Я вам на лекциях рассказывал, что как преобразуются данные в информацию. Сначала создаются справочники, потом создаются базы событий, потом выявляются причинно-следственные связи между событиями. Так вот, мы уже кое-что сделали на этом пути. Создали обучающую выборку.

Обучающая выборка представляет собой нормализованные базы исходных данных, нормализованные с помощью справочников. Отношение один к одному ко многим. Потом всё было переиндексировано. Вот. Таким образом, значит, у нас э-э выполнена формализация предметной области. Вот режим 6.4, там э-э этап формализации предметной области был первым автоматизированным этапом. До него только была был этап э-э когнитивной структуризации предметной области. Это единственный неавтоматизированный этап системно-когнитивного анализа.

6.5 Инструкции после формализации

Вот это окошко, которое сейчас появилось после этого процесса, оно сообщает студентам, учащимся, что нужно дальше делать. Мы видим, здесь написано: открыть файл исходных данных по пути. Путь полный указан этот файл. Прочитать э-э описание лабораторной работы. Ну это не нужно, потому что я рассказываю, это, во-первых, а во-вторых, у этой работы нет описания. Вот. Третье. Выполнить режимы 2.1, 2.2, 2.3.1, 3.5, 5.5, 3.4 и другие в соответствии со схемой преобразования данных в информацию, а её в знания, приведённой в режиме 6.4.

Студенты обычно так вот клацают сразу же и сидят и спрашивают: "А что нам дальше делать?" Я говорю: "Дальше делать то, что написано вот в этом окошке было. Вот это вот надо и делать дальше". "Нет, а что там было написано?" - спрашивают. Я говорю: "Ну тогда давайте опять устанавливайте работу и читайте, что там было описано". Ну, конечно, я им говорю, что там было описано. Вот там это было описано, что нужно это сделать.

Это же самое описано, ребята, и в хелпе режима 1-3. То есть если мы в этот хелп перейдём, то мы увидим э-э пункты в том же порядке, как какие режимы нам выполнять, чтобы ввести данные из внешнего источника данных, создать модели, проверить их на достоверность и решать с помощью этих моделей различные задачи.

6.6 Отображение добавленной работы в диспетчере

Вот примерно так. Здесь мы видим в диспетчере приложений, что появилось у нас приложение. Этих приложений может быть очень много в системе. Такого какого-то особого ограничения здесь нет. Единственное ограничение - ёмкость дисков, на которых, носителей, на которых расположена система.

Выходим мы из этого режима 1-3. Вот. Ну давайте сейчас я ещё раз его скриншот сделаю, чтобы показать, что приложение установлено в чате.

7. Анализ и Интерпретация Моделей

7.1 Просмотр Классификационных шкал (Режим 2.1)

И теперь, следуя вот этой инструкции, которая была в режиме 1-3, мы переходим в режим 2.1. И смотрим, что у нас там есть классификационные шкалы и градации. Тип шкал текстовый, э-э, номинальные шкалы, то есть градации неупорядоченные по смыслу. И нет какого-либо показателя или свойства, э-э такого, чтобы они были упорядочены, допустим, по возрастанию степени выраженности этого свойства, к примеру. Вот. Или, ну, обычно по возрастанию я располагаю в системе.

7.2 Просмотр Описательных шкал (Режим 2.2)

И вот есть у нас в режиме 2.2, мы видим, что у нас есть описательные шкалы и градации. Тоже они в основном текстовые, номинального типа. Но есть и порядкового типа. Размер один - это порядкового типа текстовая шкала. А размер два, я специально эту шкалу ввёл, чтобы показать пример, продемонстрировать шкалу числовую. Здесь мы видим 12 диапазонов равного размера. Числовые диапазоны.

И есть у нас шкалы, я их называю, и не только я, а так принято называть, э-э булевские или дихотомические шкалы, где у нас есть два две градации: True, False там, истина, ложь, да, нет, есть, нет. В таком плане.

7.3 Пример некорректной анкеты и важность дихотомии

Значит, я вам хочу привести пример некорректной анкеты, где есть какой-то вопрос, и если мы на него отвечаем положительно, то мы ставим птичку против этого вопроса. А если отвечаем отрицательно, то мы не ставим птичку. Можете себе вообразить такой, такую анкету, ребята? Вот, допустим, э-э согласны ли вы с тем, чтобы провести дорогу от э-э МНТК микрохирургии глаза до юбилейного микрорайона? И, значит, если вы согласны, ставите птичку против этого вопроса. Если не согласны, не ставите. Кидаете это в ящик для голосования и уходите. Потом эти бюллетени все выбрасывают оттуда, и, значит, комиссия смотрит, где там есть птички, где нет, считают, сколько было птичек, сколько за, сколько нет. Вот. И, значит, кто-то взял из этой комиссии и поставил птичку против вопроса этого. Это вполне возможно. Почему? Потому что там же э-э или ничего нет, или есть птичка. Поэтому поставить птичку можно и не заметить, что там чего-то, что там ничего не было. Вот. А чтобы это было невозможно, для этого нужно, чтобы там было два варианта ответа: согласен, не согласен. И тогда будет в любом случае птичка стоять против какого-то варианта ответа. Вот. И тогда уже так просто птичку ещё одну не поставишь. А если поставишь, бюллетень будет испорчен и не будет влиять на результаты. Ну, если там было, допустим, согласие, то степень поддержки этой этой инициативы уменьшится от того, что этот бюллетень был испорчен. То есть тоже фактически это э-э повлияет на результаты голосования. Вот. То есть можно испортить те, которые мешают э-э тому, что вы хотите. И тогда, значит, то, что вы хотите, получится, поддержано будет. Вот. Ну, это я так подсказываю вам некоторые вещи. Значит, но я сейчас хочу вам сказать, что в системе предусмотрено, что всегда есть два варианта ответа на вопрос: э-э да или нет, истина, ложь. Дихотомические э-э шкалы с двумя градациями, противоположными по смыслу. Можно сделать и больше таких градаций, что э-э абсолютно не согласен, там, не согласен, колеблюсь, непонятно, колеблюсь, э-э скорее согласен, и очень уверен, что да, что надо именно это делать. То есть сделать какие-то предусмотреть возможность каких-то степени э-э выраженности этого мнения, скажем так.

7.4 Просмотр Обучающей выборки (Режимы 2.3.1 и 2.3.2)

Теперь смотрим на обучающую выборку. Это делается в режиме 2 3 1 и в режиме 2 3 2. В режиме 2 3 1 мы видим в верхнем окне наименование наблюдения, то есть информация о том, откуда взято это наблюдение. Здесь может быть довольно большой э-э длины э-э сообщение. То есть можно особо там не не мучиться, чтобы не сокращать его. Ну, в принципе, если есть необходимость, можно длинное написать. То есть, допустим, поле номер такое-то, в таком-то хозяйстве, в таком-то году. Вот можно написать. Вот. Хотя лучше тоже покороче, потому что это всё будет в выходных формах. В левом нижнем окошке мы видим коды классов, э-э которым относятся данное наблюдение. А в правом нижнем окошке - коды признаков соответствующих или значений факторов, или значений свойств этого наблюдения. Вот. И есть ещё один вариант отображения обучающей выборки, но он не всегда возможен, а только тогда, когда использовался интерфейс 2 3 2 2 для ввода э-э обучающей, ну, для, для формализации предметной области. Значит, заходим в режим, то есть не в режим, а в папочку заходим, где у нас исходные данные. Открываем файл исходных данных и смотрим, какой у нас файл исходных данных и какое у нас файл обучающей выборки. Вот. Значит, они совершенно одинаковые по структуре. Вот. Те же самые поля. Э-э, ну, размер, конечно, у них э-э другой, но, в общем, те же самые поля. Мы видим, что э-э есть наименование объекта наблюдения совпадает полностью. Видим э-э две колоночки классификационных шкал на жёлтом фоне и колоночки описательных шкал на белом фоне. И мы видим, что в Экселе там пронумерованы колоночки A, B, C, D, E, F и так далее, латинский алфавит. А в системе Эйдос они пронумерованы 1 2 3 4 5. И вот мы видим, что они совпадают, и мы видим, что э-э первая колоночка не является шкалой, а является наименованиями объектов наблюдений. А вот вторая и третья мы задавали только что в экранной форме, что это диапазон классификационных шкал, номера колоночек. А с четвёртой по двенадцатую - это колоночки описательных шкал. Всё это вот мы видим здесь в этой форме. Вот сейчас я эту форму скриншотик помещу в чат Teamса.

7.5 Обзор Матриц моделей (Режим 5.5)

Вот таким образом у нас э-э проведена формализация предметной области. И сделано всё, что необходимо для того, чтобы начать создавать модели и идти дальше по этапам системно-когнитивного анализа. Ещё я хочу ваше внимание обратить, что вот э-э в файле исходных данных, видите, э-э указаны э-э градации шкалы размер один, э-э как порядковые шкалы: три, четыре, три, два, четыре там, один, четыре. Вот, видите? То есть указано, сколько всего интервальных значений и каким конкретно является данное интервальное значение: минимальным из четырёх, средним из четырёх, максимальным из четырёх. А в числовой шкале это делается автоматически: ищется минимальное, максимальное значение и делится на заданное число интервалов.

Ещё хочу ваше внимание обратить на то, что в числовых шкалах всегда надо указывать единицы измерения. Обязательно.

Вот. Теперь мы можем перейти к следующему этапу. Это синтез и верификация моделей. Что такое синтез модели, ребята? Кто может догадаться? Ну это, наверное, не очень сложно догадаться. Это создание модели. Вот. А верификация - что такое? Верификация - это проверка их на достоверность, проверка достоверности модели.

Значит, вот у нас, э-э, мы видим в этой экранной форме, создаётся три статистических модели, семь системно-когнитивных моделей, модели знаний. Мы видим это и в режимах, э-э, 6-4 перечислены, какие модели создаются. Вот, и в описаниях, э-э, в хелпах различных режимов.

Значит, мы э-э задаём здесь э-э также, на каком процессоре выполнять расчёты: на центральном или на графическом. На графическом мы можем выполнять расчёты только в случае, если у нас видеокарта поддерживает OpenGL, и это обычно для видеокарт Nvidia характерно, GTX. Значит, если на графическом процессоре, то мы можем либо визуализировать промежуточные стадию исполнения, либо нет процесса исполнения. На центральном процессоре всегда визуализируется. И может быть, э-э, использовано два способа расчёта. Э-э, для количества информации там всегда одинаково считается, а для э-э другого интегрального критерия может быть упрощённый вариант или полный. Результаты при этом мало чем отличаются. Упрощённый вариант, он работает значительно быстрее. Мы можем э-э только создавать модели, только верифицировать модели, или можем создавать или и верифицировать. Если у нас очень много объектов наблюдения, и мы хотим на основе такой обучающей выборки э-э проверить достоверность модели, то у нас может затянуться этот процесс верификации. Тогда лучше сделать отдельно синтез модели, он довольно быстро проходит, и потом верификацию провести на центральном или на графическом процессе, достаточно быстро. Вот верификация, она значительно дольше осуществляется, потому что каждый объект обучающей выборки, каждое наблюдение сравнивается со всеми классами. Это делается в каждой модели. Вот. Если у нас объектов наблюдения очень много, признаков очень много, классов много, то там большой объём расчётов проходит. Поэтому делается так: значит, только синтез модели, а потом верификация на небольшом количестве. Для этого используется бутстреп-подход. То есть часть э-э обучающей выборки используется для проверки модели на достоверность. Эта часть может быть небольшой, ну, скажем, 10% там, например, или 5%. Если большая очень обучающая выборка, то этого вполне достаточно, чтобы оценить достоверность модели. Вот это подмножество для того, чтобы оценить достоверность модели, может определяться различными способами, которые здесь несколько в системе Эйдос. Но некоторые работают очень быстро, некоторые помедленнее работают. Вот. Быстрее всего работают, когда объекты с какого-то по какой-то используются для проверки модели на достоверность.

Значит, ну сейчас мы используем параметры по умолчанию. Значит, здесь также указано, ребята, что вот эта модель, которая сейчас будет создаваться, она составляет 60% от того, что позволяет версия системы Эйдос непрофессиональная. И то она, видите, имеет довольно большие возможности. Но в ней есть режимы, э-э позволяющие ещё больше объёма информации обрабатывать. Но где тогда надо здесь вот задавать э-э использовать не все результаты распознавания, а только наиболее достоверные. Ну, например, вместо 10% взять 5% или 1%. Тогда можно обрабатывать огромные базы данных, вот, просто огромные, и всё работает нормально.

Сейчас мы по умолчанию будем всё это делать. Просто нажимаем Окей. Э-э, на такой маленькой модели довольно быстро всё это происходит, синтез модели и проверка их на достоверность. Сначала происходит синтез модели, потом проверка их на достоверность. Путём решения задачи идентификации объектов обучающей выборки.

Значит, э-э, можно это делать с помощью решения других задач. Можно, допустим, путём решения задачи прогнозирования решать, оценивать, насколько модели достоверны, или путём решения задачи принятия решений, управления. Вот если мы, допустим, решим задачу прогнозирования, то нам надо будет подождать немножко, там годик-два, и мы увидим, достоверная модель или нет, осуществились ли вот эти прогнозы, которые дала нам система на основе моделей. Ну это я так подшучиваю немножко. Можно сделать прогнозирование на ретроспективных данных. То есть можно взять данные конца двадцать первого года и спрогнозировать, что будет в начале двадцать второго года. Совпало, не совпало, мы можем оценить достоверность модели путём прогнозирования. Можем? Да, можем. Но это логически сложнее, чем путём решения задачи идентификации, потому что задача прогнозирования вообще сама по себе сложнее, чем задача идентификации.

И можем также проверять достоверность модели путём решения задачи принятия решений. Взяли, выработали какие-то решения, э-э реализовали их на практике. Вот. И у нас э-э получилось там лучше или хуже тот или иной результат. Вот. И мы видим, что если у нас объект моделирования или объект управления перешёл в целевое состояние, тогда, значит, у нас модель управления адекватная, модели достоверная. Не перешёл, значит, она неадекватная. Вот. Но я могу вам сказать, что вот это вот э-э решать, оценивать достоверность модели путём решения задачи принятия решений - это вообще-то э-э авантюра, опасная очень, потому что если модель недостоверная, профанация, я бы сказал даже, научных подходов, то, значит, если модель недостоверная, могут получиться результаты совершенно нежелательные. Ну, например, мы берём, э-э оцениваем э-э э-э состояние предприятия, создаём модели на основе этих оценок, и потом э-э рекомендации вырабатываем, что нужно сделать на данном предприятии, чтобы оно отдалилось от состояния банкротства. И мы даём эти рекомендации, это всё выполняется руководством предприятия, и оно обанкрочивается. Тогда нам начинают нас искать с собаками, грубо говоря, и спрашивать: "Что это вы там нам посоветовали? Мы потерпели убытки там вот такие-то, такие-то, вы в этом виноваты, давайте там, ну, в общем, претензии могут предъявить". Значит, э-э что отсюда следует? Что решение задачи э-э принятия решений является плохим способом для оценки достоверности модели, опасным, понимаете? Нельзя вырабатывать рекомендации какие-то и реализовывать их на практике, используя для этого непроверенную модель, неизвестной достоверности. А вдруг модель недостоверна, и эти вот рекомендации будут неадекватны, и результат будет совершенно нежелательный, не тот, который хотелось бы. И вот на защитах я, ребята, встречал случаи, когда соискатель степени описывает модель, и я жду, ну вот сейчас вот скажет, что мы эту модель проверили на достоверность, она оказалась достоверной, и мы выработали вот такие-то решения, такие-то рекомендации, их реализовали и получился вот такой-то дополнительный эффект. И вдруг этот соискатель говорит: "Вот мы сделали вот такие-то модели, разработали. Выработали на основе них рекомендации, на основе этих моделей. И получился вот такой положительный эффект". Я говорю: "Так, стоп. Подождите, секундочку. Вы вы разработали модели и сразу выработали рекомендации и сразу их применили на практике, что ли? А достоверность модели вы что, не измеряли, что ли?" Они говорят: "Нет". Я говорю: "Ну тогда это ж вообще э-э очень опасное вы совершили действие. Ведь ваши модели могла оказаться недостоверной. И ваши рекомендации могли оказаться не полезными, а вредными для предприятия. Как можно в таких условиях, не зная достоверность модели, э-э давать рекомендации? Это крайне рискованно, опасно и вообще это, в общем-то, я сказал уже, профанация и авантюра". Вот. И тут соискатель отвечает смелой улыбочкой, э-э девушка: "А у нас всё получилось хорошо. У нас э-э эффект получился положительный". Я говорю: "Ну тогда, значит, вам просто очень повезло, понимаете? А могло не повезти. Могло получиться совершенно иначе. И что тогда было бы?" Вот. Ну, слава богу, я я понимаю так вот, осознаю, что на самом деле ничего там никаких рекомендаций никто не применял на практике, потому что страшно применять эти рекомендации на практике. Кроме того, э-э руководитель, он, видимо, человек, так сказать, опытный, э-э и не зря он руководитель, и уже много лет успешно руководит этим предприятием. Если бы мы дали какие-то бредовые рекомендации, то он бы подумал бы, на всякий случай, не буду я их э-э реализовывать эти рекомендации. Нужно сначала надо убедиться, что это всё разумно, а потом уже использовать для управления.

7.6 Обзор полученных моделей (статистических и системно-когнитивных)

Ребята, значит, поднимите руку, у кого вот эта картинка на экране, которая у нас получилась в результате синтеза моделей. Ну, естественно. Так. И Анастасия. Ну, молодцы. Вообще прямо вы меня поражаете. Всё, можете руку опускать.

Следующий этап у нас - это мы смотрим на сами эти модели. Мы это делаем в режиме 5.5. Смотрим на сами эти модели, которые у нас получились.

У нас есть модели три статистических модели: матрица сопряжённости, матрица условных безусловных процентных распределений и матрица системно-когнитивных моделей. Я на лекции вам показывал, как рассчитываются эти модели, и скриншотики помещал э-э в чат лекции. Вот. И там мы видели, что у нас используются э-э частные критерии, и получаются 10 у нас моделей: три модели статистических и семь моделей системно-когнитивных. Они сходны друг с другом по смыслу, потому что они позволяют сравнивать фактические и теоретические абсолютные частоты и э-э условные и безусловные относительные частоты встречи тех или иных значений факторов или значений свойств у объектов различных категорий.

И вот так получается у нас 10 вот таких вот матриц. Такие матрицы, я вам не упоминал на лекции, но это можно сейчас упомяну, являются определителями Вронского. То есть это матрицы, в которых в ячейках находятся формулы, по которым приводятся расчёты значений этих ячеек. Вот эти формулы здесь приведены. То есть это определители Вронского. То есть можно сказать так: в системе Эйдос используются определители Вронского для создания модели.

Значит, смотрим модель частот, абсолютных частот. Вот эта модель. Смотрим в самую нижнюю строчку. Вот. И в самый правый краешек смещаемся. Вот что мы здесь видим, ребята? Мы видим, что у нас э-э здесь разные количество встреч тех или иных значений свойств у объектов различных категорий. Ну я привожу всегда вот такой пример, уже выработал, что рассказывать. Вот смотрим мы размер под руку. Вот у нас три раза он встретился этот размер под руку у средств связи и три раза он встретился у элементов компьютера. Ясно, да? Как вы думаете, эта тройка и эта тройка, они имеют одинаковый смысл или разный? И не только в данном случае, а, как говорят, вообще говоря, в общем случае. Как вы считаете? Вот здесь, сейчас, ребята, вы пока там думаете, я вам задам такой вопросик, похожий на анекдот, анекдотичный вопросик. Вот э-э посчитайте, пожалуйста, значит, э-э у нас два крокодила летит прямо, два крокодила прямо летят, а три крокодила у нас зелёного цвета. Сколько всего крокодилов? Или лучше всего, сколько всего верблюдов спросить? Что скажете? У нас нет верблюдов. А? У нас нет верблюдов. Ну да, нет верблюдов. Вот. И, значит, э-э крокодилы вообще-то э-э и те, которые летят прямо, и которые направо, э-э таких не существует, вы же знаете, да, про это? Ну только разве что с на за хвост нас мотите, раскрутите и кинуть, да? Или с самолёта кинуть. Так вот, э-э к чему я так клоню? К тому, что у нас э-э вот в этих колоночках э-э элементы компьютера, средства связи, у нас, вообще говоря, разные э-э смысл этих колоночек и разное число наблюдений объектов, относящихся к этим колоночкам. То есть, когда мы складываем, допустим, э-э ну, допустим, 1/3 э-э прибавляем 1/5. То как мы это делаем? Мы берём и переводим в одну единицу измерения, переводим к общему знаменателю, правильно? Когда это всё будет в одной единице измерения, тогда можно складывать. Понимаете? Или, допустим, там 1,2 кубо- там кубометра прибавить, допустим, сколько-то там литров. Можно прибавить? Можно, только надо перевести в одну единицу измерения, а потом прибавлять. Согласны, да? Сначала неправильную дробь перевести в правильную, потом, значит, перевести к одной единице измерения, к общему знаменателю. Вот, одной единице измерения, к общему знаменателю, либо складывать. Так вот, э-э то же самое касается и здесь этих моделей. Мы видим, что у нас э-э средств связи три наблюдения было по всей выборке, а элементов компьютера семь. Поэтому можем определённо сказать, что вот эти вот троечка, здесь троечка и здесь - это вообще-то совершенно разные троечки. Их нельзя сравнивать, нельзя говорить, что вот это вот, то, что здесь три и здесь три - это означает, что вот этот признак под руку имеет одинаковое значение или одинаково характерен для объектов этих категорий. Потому что это абсолютное значение. Нам нужно перейти к относительным значениям, чтобы мы могли это обоснованно сказать.

7.7 Матрица условных и безусловных частот

И вот мы это делаем. Переходим к относительным значениям. Вот. И смотрим. Размер под руку. Вот, у нас получается 100% средств связи э-э обладают этим признаком, и почти 43% элементов компьютера тоже обладают этим признаком. То есть мы перевели эти значения к относительным величинам, от абсолютных перешли к относительным. То есть, если вот взять хелп этого режима, то мы увидели, что от матрицы абсолютных частот, от матрицы сопряжённости, мы перешли к матрицам условных, безусловных процентных распределений.

7.8 Матрицы системно-когнитивных моделей (Инф1, Инф3 - ХИ-квадрат)

Теперь мы можем сравнивать уже эти вот э-э э-э вероятности, относительные, ну, вернее, относительные частоты встречи этих признаков. Они стремятся к вероятностям при неограниченном увеличении объёма выборки. И мы видим, что этот признак намного более характерен для средств связи, чем для элементов компьютера. Но и для тех, и для тех он характерен больше, чем вообще в среднем для объектов выборки, которых только 30% с этим признаком. Ясно, да?

Теперь возникает такой вопрос, ребята. Как вот это всё сравнивать? Если у нас матрица вот такая малюсенькая, когда вот у нас там всего 19 этих э-э объектов, 49 там признаков, то можно ещё как-то пытаться глазами вот так вот сравнивать и понимать, пытаться понять смысл этого. Но даже если мы этот смысл поймём, то мы его не сможем использовать при решении задач, потому что э-э для того, чтобы использовать его для решения задач, для этого надо, чтобы это было сравнение проведено в электронной форме, в самой системе. И находились результаты этого сравнения в базах данных. И система могла из этих баз данных брать это результаты этого сравнения и решать на основе этого задачи. То есть ясно, что нужно это автоматизировать сравнение. И оно и реализовано в системе. Вот в этой помощи как раз и показано, как это сделано. Вот, показаны частные критерии, как это сравнение проведено, и как получены матрицы э-э системно-когнитивных моделей, в которых уже результаты сравнения приведены вот этим матрицам.

Смотрим на модель Инф1. Эта модель чем интересна? Здесь используется количество информации Харкевича для расчётов. И мы видим в битах, ребята, я хочу подчеркнуть, прямо нормированное в битах, битом, всем известным, какое количество информации мы получаем о том, что объект относится к тому или иному классу, средством связи или к элементам компьютера, если узнаём, что у него есть такой признак размер под руку. И мы видим, что мы получаем 0,779 тысячных бита там семью знаками расчёт, но показано три с округлением. Мы получаем о том, что это средство связи и 0,231 тысячная бита о том, что это у нас элемент компьютера. То есть у нас уже здесь указано количество информации. Это уже результат сравнения. И его можно использовать прямо вот этот результат для решения задачи идентификации, прогнозирования, принятия решений и исследования моделирования предметной области, что мы рассмотрим подробнее на следующем занятии.

А сейчас посмотрим ещё на модель Инф3. Модель Инф3 - это модель хи-квадрат. И смотрим тоже, во-первых, у неё очень интересная особенность есть, э-э замечательная особенность, э-э за которую э-э Пирсону, который предложил хи-квадрат, просто, ну как это сейчас говорят, респект. Просто уважение можно выразить ему, восхищение. Значит, сумма сумма по колонке равна нулю, и средняя по колонке равна нулю, ребята. И то же самое касается строк: сумма по строке равна нулю, и средняя по строке равно нулю. То есть эта модель прекрасно сбалансирована, прекрасно сбалансирована. И мы видим, что для э-э признака размер под руку, значение элемента хи-квадрат, одного из слагаемых, там потом по всей матрице суммируется. Э-э для средств связи 2,1 значение в степени характерности, а для элементов компьютера 0,9. И и для средств связи, и для элемента компьютера фактическое значение значительно превосходит теоретическое значение и указывает на сильную связь между наличием этого признака и принадлежностью объектов к этим классам. А вот других вот колоночках мы видим минусы, видите? Что это не стул, не телефон, или что там у нас? Нет, это это телефон, но это не аксессуар, не мебель, не спорт-инвентарь. Видите, да? Вот эта строчка показывает.

7.9 Аналогия с Магическим квадратом

Я спрашиваю, хочу вас спросить, вы слышали, что такое э-э магический квадрат? Кто-нибудь слышал? А, ребят? Это квадратная матрица, прежде всего. Квадратная. В которой э-э сумма строк, сумма значений в строке э-э во всех строках одинаковая. И сумма значений в колоночке тоже одинаковая. И такая же, как в строке. И больше того, по диагонали тоже такая же. Ну, правда, это значение может быть различным. Ну вот, например, магический квадрат, у котором 15 это значение, сумма элементов строки, сумма элементов колоночки и диагонали. А вот э-э модель хи-квадрат э-э э-э в системе Эйдос рассчитывается, очень напоминает магический квадрат, только значение ноль и в сумме колоночки, и в сумме по строкам. И, что интересно, и среднее тоже ноль. А в магических квадратах среднее не ноль, не не равно этому числу, э-э сумме которых равны значения этих элементов. Но есть ассоциация, есть как бы аналогия какая-то. То есть это интересно. Если бы в системе Эйдос э-э число классов и число признаков было одинаковым, то тогда модель хи-квадрат была бы магическим квадратом. Но сейчас вот у нас мы видим, классов 14, а признаков 49. Тем не менее, это какой-то вариант э-э похожий, что-то похожего на магический квадрат. Я думаю, это интересно, прикольно, я бы сказал.

8. Заключение и Планы на будущее

8.1 Важность проверки достоверности моделей

Вот. И надо сказать ещё пару слов о матрице хи-квадрат. Значит, э-э я вам уже упоминал, что если мы не измерили достоверность моделей, то решать задачи нельзя. Если вы услышали, что кто-то сделал какую-то модель и начал применять её для решения задач, то у вас сразу должно возникнуть в памяти то, что я вам рассказывал. Вы должны сразу же вспомнить и спросить: "Скажите, пожалуйста, а вот эта модель, которую вы применяете для идентификации, прогнозирования, принятия решений, исследования, вы её достоверность проверяли эту модель или нет?" И если вам скажут, что нет, то это значит тогда вообще дальше вообще говорить не о чем. Что сначала нужно обязательно проверить достоверность, а потом уже решать задачи. Ну, на меня ссылаться не обязательно, это не только я говорил, это все говорят, что сначала надо достоверность, а потом это. Но почему-то многие этого не делают и вообще забывают про это. Вы должны помнить. Я вам сейчас вот напоминаю, что как только вы слово модель услышите, сразу вы должны вспоминать про то, что я говорил, что нельзя модели применять, пока не измерили достоверность. Если достоверность достаточно высокая, тогда можно применять модель на практике.

8.2 План следующих занятий

Так вот, э-э в системе Эйдос э-э предусмотрено измерение достоверности модели в режиме 3-4. Вот режим 3-4, смотрите, видите, показано. И это делается путём решения задачи идентификации. Я вам сказал, что задача прогнозирования сложная является для этой цели, э-э задача принятия решений вообще является халтурой, авантюрой и профанацией. Вот. А задача идентификации, она достаточно проста, и э-э путём идентификации объектов обучающей выборки вполне можно оценивать достоверность модели безопасно, не принимая никаких решений, которые бы исполнялись бы на практике, а просто для самого себя узнать. Вот мы создали обобщённые образы классов, а система в состоянии э-э идентифицировать объекты, относить их к классам правильно или нет? Если она это делать не может, то тогда вообще говорить не о чем. То есть можно так немножко слукситься, расстроиться и понять, что надо тогда переделывать модель вообще-то. Какие-то другие варианты искать, другое число диапазонов числовых, использовать адаптивные интервалы, какие-то ещё добавить шкалы описательные. В общем, что-то нужно делать. Нельзя этой моделью пользоваться, ребята, если она имеет низкую достоверность.

Вот. И поэтому мы и смотрим, как это делается в системе Эйдос. Следующий наш пункт нашего этого рассмотрения - это оценка достоверности моделей. Как это делается в системе? Значит, есть режим 3-4, есть ещё другие режимы, где тоже эта же информация выводится в более подробном виде. Вот главная экранная форма этого режима. И есть ещё у него два хелпа. Хелп по мерам достоверности, количественные меры достоверности модели. И э-э хелп по частотным распределениям э-э количество истинных и ложных, положительных и отрицательных решений при различных э-э критериях достоверности, которые применяются в системе Эйдос. Я вам привожу оба этих хелпа.

И теперь начинаю вам рассказывать, что мы здесь видим. Смотрите, э-э каждая модель два раза представлена с разными интегральными критериями. То есть я вам буду вынужден буду одновременно рассказывать и про меры достоверности, и про интегральные критерии. Мы видим, что модель Inf3 ХИ квадрат, э-э есть два интегральных критерия, как и в других моделях - это сумма знаний и семантический резонанс знаний. И по этим интегральным критериям, вернее, с их использованием получается по критерию достоверности F-мера Ван Ризбергена 0,920 достоверность модели. Что такое 0,920? Максимальное значение - единица. А по критерию L1, который я предложил, 0,976 достоверность. А по критерию, значит, L2, которое тоже я предложил, 0,929. Значит, ну, достоверность э-э 0,976 - это высочайший уровень достоверности, ребята, потому что максимальное значение равно единице. Этот критерий достоверности, он э-э нормирован и имеет значение от нуля до единицы.

И вот э-э давайте, сейчас я начну вам рассказывать, и уже у нас осталось не так много времени занятия. И на следующем занятии вы мне напомните, что мы начнём с рассмотрения критериев достоверности моделей.

Ребята, значит, у нас на этом конец занятия. На следующем занятии, пожалуйста, напомните мне, что мы начинаем с этого места, э-э то есть мы начинаем э-э практическое занятие с того, что быстренько восстанавливаем лабораторную работу вот эту, создаём модели, и потом я вам показываю вам дальше э-э про оценку достоверности модели, всё рассказываю. Потом мы выберем наиболее достоверную модель, и в ней будем решать задачи. Это будет ещё одно занятие. После этого мы на каждую э-э лабораторную работу будем тратить одно занятие. На это мы потратим три на первую. Вот. А на следующие мы по одному будем тратить: обработку текстов, обработку изображений, будем быстро проходить.

Ну всего самого хорошего вам. До свидания. Здоровья и успехов.