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

***Российская Федерация***

***Потемин Илья ПИ2103***

**119 Математическое моделирование и анализ данных в садоводстве. Лабораторная 2. Исследование моделируемой предметной области 2020-10-06**

Заголовок: Анализ данных в садоводстве: Дендрограммы, конструкты и система поддержки решений "Эйдос"

Резюме:

1. Введение и контекст

Лекция посвящена лабораторной работе №2 по дисциплине "Математическое моделирование и анализ данных в садоводстве" для группы ПВ-2041. Рассматривается вопрос исследования и моделирования предметной области садоводства с использованием интеллектуальной системы "Эйдос".

2. Анализ совместимости состояний с помощью дендрограмм

Представлен анализ дендрограммы для определения совместимости будущих состояний объекта моделирования. Состояния, находящиеся в одном кластере, считаются совместимыми, так как их системы детерминации (обуславливающие факторы) схожи. Достижение состояний из разных кластеров, особенно на высоком уровне различия (противоположных полюсах конструкта), затруднительно или невозможно одновременно. Это важно для корректной постановки целей управления.

3. Ограничения моделирования и роль опыта

Подчеркивается, что модель строится на основе прошлого опыта (данных). Однако опыт не гарантирует идентичного поведения системы в будущем из-за изменений технологий, сортов и т.д. Приводятся примеры выдающихся личностей (Мичурин, Тесла, Пустовойт, Лукьяненко), чьи достижения сложно повторить, так как их методы могли выходить за рамки общепринятых "конструктов" или основываться на интуиции и будущих научных знаниях.

4. Концепция конструктов

Вводится понятие "конструкт" – биполярное понятие (например, близко-далеко, низко-высоко, тепло-холодно), имеющее спектр промежуточных значений. Мировоззрение человека определяется его системой конструктов. Обучение и развитие расширяют эту систему, добавляя новые конструкты или увеличивая их размерность и диапазон. Пространственно-временные конструкты легли в основу точных наук, но существуют и другие (цвет, вес, возраст, оценки и т.д.).

5. Анализ влияющих факторов

Обсуждается расчет матрицы сходства факторов, влияющих на объект моделирования. Визуализация с помощью круговой диаграммы и графика межкластерных расстояний показывает, какие факторы оказывают сходное, а какие – противоположное влияние на результаты (например, финансовые и натуральные показатели выращивания).

6. Связь факторов и состояний

Выдвигается гипотеза, что группы факторов со сходным влиянием (один кластер факторов) обуславливают переход объекта в состояния одного кластера состояний (один полюс конструкта), а противоположные группы факторов – в состояния другого кластера (противоположный полюс). Таким образом, системы факторов и системы состояний взаимосвязаны через конструкты.

7. Система поддержки принятия решений "Эйдос"

Представлен развитый алгоритм принятия решений в системе "Эйдос", который превосходит упрощенный SWOT-анализ. Он позволяет:

\* Работать с несколькими целевыми состояниями одновременно.

\* Оценивать корректность постановки целей (их совместимость).

\* Учитывать ограничения (например, невозможность использования определенных факторов по финансовым или технологическим причинам).

\* Прогнозировать результаты при измененной системе факторов.

\* Заменять недоступные факторы на другие со сходным влиянием.

\* Объективно рассчитывать значимость факторов (в отличие от экспертных оценок в SWOT).

\* Адаптироваться к новым данным, обновляя модель.

8. Технические аспекты и обсуждение

В ходе лекции происходит обсуждение технических проблем с запуском системы "Эйдос" у студентов, даются ссылки на научную статью и свидетельство Роспатента, связанные с используемыми методами кластеризации.

Детальная расшифровка текста:

1. Введение и контекст

Здравствуйте, ребята.

Здравствуйте. Здравствуйте, Евгений Вениаминович.

Сколько лет, сколько зим.

Сегодня у нас 6 октября 2020 года. Вторая пара, 9:45-11:15. Лабораторная работа номер два с группой ПВ-2041 по дисциплине Математическое моделирование и анализ данных в садоводстве.

Мы сейчас рассматриваем вопрос исследования моделируемой предметной области. Занятие ведёт профессор Луценко Евгений Вениаминович.

2. Анализ совместимости состояний с помощью дендрограмм

Вот.

Таким образом, давайте посмотрим дендрограмму. И по этой дендрограмме я вам дам комментарии.

Вот, мы здесь видим, какие будущие состояния объекта моделирования совместимы по системе обуславливающих их факторов, а какие несовместимы. Если руководство ставит задачи как целевые состояния достичь состояний, которые находятся в одном кластере, то это вполне возможно, потому что это означает, что если они в одном кластере, что их система детерминации, которая обуславливает переход объекта моделирования в эти состояния, она очень сходна.

Если же различия выше вот здесь вот уже, тогда, возможно, есть некоторая проблема в этом. А если они вообще находятся эти состояния в противоположных кластерах по самому высокому уровню сходства, то есть уже сильно отличающиеся друг от друга, на различных полюсах конструкта, тогда это одновременно недостижимо.

Вот. То есть если целевое состояние находится в одном полюсе конструкта, а другое целевое состояние в другом полюсе конструкта, то тогда это альтернативные состояния, несовместимые, одновременно недостижимые.

И мы это видим очень наглядно. То есть это важно для того, чтобы определить корректность поставленных перед нами целей.

3. Ограничения моделирования и роль опыта

По, значит, надо отметить, что эти вот формы, они отражают весь опыт предыдущий выращивания этих сельхозкультур. То есть если мы ввели эти данные и получили такие результаты, это означает, что по нашему опыту это вот так.

Но наш опыт не является, так сказать, полной гарантией, что в будущем тоже будет так, потому что всё меняется, технологии развиваются, сорта новые появляются, технологии новые появляются. И возможно, что в будущем станет возможным и то, что сейчас мы нам кажется невозможным.

Ну, жизнь показывает, что это происходит, скажем так. Но каким образом это получается, никто не поймёт. То есть какие-то люди появляются очень высокой степени креативности, компетентности, типа Пустовойта, Лукьяненко, и у них получается то, что у других раньше не получалось. Ну, Мичурина того же. Вот, чудеса какие-то люди творят. То есть, причём в буквальном смысле попахивает чудесами, потому что я вот спрашивал наших профессоров, вот сейчас прямо, ну как сейчас, год назад, может быть, там, два года назад. Как они относятся к Мичурину? Они говорят: "Ну он просто вот волшебник, понимаете?" То есть его повторить не удаётся. Почему? Его методы? Потому что вообще непонятно, как он этого добивался, того, что он добился.

То есть с точки зрения современной науки это не совсем понятно, я бы сказал, и вообще непонятно. Ну это примерно как Тесла. Самородки такие, самородки, понимаете? Какие-то выдающиеся люди, которые что-то там как-то интуитивно понимают, а может быть, даже и логически понимают. Но мы не можем их, перенять их уровень.

Вот. Вот то, что делал Тесла, это сейчас вся наша цивилизация на этом основана практически, на его изобретениях. Абсолютно непонятно, откуда он всё это брал. Когда его просили рассказать, он рассказывал, ничего было непонятно, что он рассказывал. Он пытался там лекции читать, ничего никто никто ничего не мог понять, что он рассказывает.

То есть сильно очень отличался он по своему уровню и по характеру понимания этих вопросов от окружающих. Сейчас вот с помощью, да, он не пользовался достижениями науки своего времени. То есть когда пытаются методами науки понять, что он там делал, то не очень получается, понимаете?

Вот. А у него-то это работало всё. А каким образом он это делал? Непонятно. То есть получается, что он основывался на науке будущего, которая нам ещё неизвестна.

3.1. График межкластерных расстояний

Это график увеличения межкластерных расстояний. Мы здесь видим, что сначала объединяются будущие состояния в кластеры очень хорошо, на минимальном уровне различия. Ну, то есть фактически это значит, что они вообще не отличаются по системе детерминации. Потом довольно неплохо это получается. И начиная с какого-то момента, где-то уже ближе к концу, различие резко возрастает между кластерами. Вот этот график мы видим. Когда происходит объединение вот где-то в области 40%, то потом наблюдается резкий рост расстояния между кластерами. То есть дальше можно сказать так: если здесь сходные объединяются кластеры, здесь уже объединяются кластеры, э-э, классы, которые сильно отличаются друг от друга, но всё-таки они объединяются на высоком уровне различия.

3.2. Ссылки на материалы

Вот. И здесь вот ссылка на статью есть, которая посвящена этой тематике, кластеризации. Сейчас я вам эту ссылочку в чат отправлю, чтобы вы могли познакомиться.

И есть ещё ссылка на свидетельство Роспатента на этот режим кластеризации. Вот. Когда я его сделал.

Эту статью, ребята, я вам советую почитать. Здесь даётся критика традиционных классических методов кластерного анализа. И потом предлагается метод кластерного анализа, который решает эти проблемы, которые есть у методов, которые сейчас применяются классических.

Вот, всё здесь приводится это, расписывается.

4. Анализ влияющих факторов

Значит, теперь дальше двигаемся. Смотрим, насколько сходны друг с другом различные факторы, оказывающие влияние на объект моделирования. Для этого мы рассчитываем матрицу сходства.

4.1. Технические проблемы и обсуждение с аудиторией

Ребята, у кого-нибудь из вас получается делать то, что я рассказываю на своих компьютерах?

М?

У кого-нибудь из вас на своих компьютерах получается делать то, что я рассказываю, или нет?

На прошлом занятии, на этом почему-то Эйдос не хочет запускать программу. Ну то есть я захожу на 1.3, запускаю скачать с облака, и он выводит ошибку постоянно. В прошлый раз такого не было.

Сейчас попробуйте, может получится. Это связано с тем, что одновременно обращение много людей, ну, многих людей обращение происходит. У меня тоже сейчас два раза не получилось, а потом получилось. То есть когда вы пытались, у меня не получалось. Потом вы перестали...

Нет, я имею в виду, я имею в виду, захожу, вот нажимаю скачать с облака, он мне не хочет открыть, само облако не открывает. Он мне пишет сразу ошибка Эйдос и всё.

Да? Ну посмотрю. В принципе, я могу и сейчас посмотреть, что там такое.

А вот сейчас попробуйте.

Нет, всё равно убивает ошибка base 1024.

Хм.

Да. Сложно это всё дело. То есть даже когда нажимаешь просто на кнопочку, да, вот эту, то уже всё вылетает, да? Вот скачать приложение из облака нажимаете, и всё, сразу ошибка, да?

Да, да, да, сразу же.

У меня вот сейчас вот нормально загрузилось, вы ж видите, да? Или я не сделал демонстрацию экрана? Сейчас я поделюсь.

А, у вас есть совместный доступ. То сейчас вы должны видеть, что у меня открылось. Ну то есть оно работает вообще-то.

Может быть из-за того, что сама первая программа не запускается, которая Eidos Start Eidos?

Да.

Ну Start Eidos не запускайте, запускайте прямо сам исполнимый модуль.

Ну я его запустил, вот получается такое.

Да?

Ой, боже мой. Ну когда на одном компьютере работает, на другом нет, то получается, что проблема локализуется в компьютере. То есть это не проблема FTP сервера, программы. Понимаете? Если она у меня работает, а у вас нет, то это что-то связано с вашим компьютером. Попробуйте его перезагрузить, не знаю, заново систему скачать. Должна быть 3.10.2020 должна быть версия. Вот. Установить заново её полностью.

4.2. Визуализация сходства факторов

Вот. Теперь давайте смотреть, как факторы влияют. Значит, сейчас мы посчитали матрицу сходства факторов и можем посмотреть, как они влияют в виде круговой диаграммы. При этом мы тоже получим круговую диаграмму, на которой ничего непонятно. Вот. Потому что очень большое число факторов. Ну, по крайней мере, так сравнительно большое, достаточно большое, чтобы было непонятно, вот так скажем. Вот. Ну поэтому нужно применить какие-то параметры вот здесь, которые сделали б более понятно, то есть информация, чтобы было меньше представлено.

Ну это, по-моему, я переборщил. Надо давайте здесь вот сделаем 40. Вот. И посмотрим график. Он становится более разумным.

5. Связь факторов и состояний

Значит, что мы здесь видим, ребят? Значит, вот слушайте внимательно, что я сейчас скажу. Сейчас я интересную вещь скажу. Что вот эти вот факторы, которые вверху объединены красными линиями, они оказывают сходное влияние на результаты выращивания финансово-экономические и натурального выражения. То есть результаты выращивания в стоимостном и натуральном выражении. И вот эти факторы, которые внизу находятся, тоже объединены красными линиями, и тоже оказывают сходное влияние на результаты выращивания в стоимостном и натуральном выражении. А вот эти две группы факторов, верхняя и нижняя, оказывают противоположное влияние.

И можно что здесь сказать? Что мы видели кластеры классов, отражающих будущие, соответствующие будущим состояниям. И видели, что кластеры эти тоже делятся на два, на две группы, образующие полюса конструкта. И я так думаю, что одна группа факторов обуславливает переход в группу кластеров, которые на одном полюсе конструкта, а другая, которая в группу кластеров, которая на другом полюсе конструкта. То есть это противоположные системы факторов влияют так, что объект переходит в противоположные состояния.

Ну это как раз и есть факторы, это ж и есть система детерминации. Получается, что эти вот две системы детерминации, вот эта вот и вот эта, они несовместимы. Что мы делаем либо это, либо это, и объект моделирования переходит либо в состояние, соответствующее одному полюсу конструкта, либо другому.

6. Концепция конструктов

6.1. Необходимость объяснения

Ну здесь нужно, видимо, вам объяснить, что такое конструкт. Ребят, вы я вам объяснял, что такое конструкт, нет? Признавайтесь. Кто-нибудь помнит?

Нет.

Не было, да?

6.2. Определение и примеры

Значит, ну это очень важное понятие, оно из области психологии, но оно гораздо шире имеет значение, чем только для психологии. Вот, скажем, для нас оно тоже имеет большое значение. Ну вы помните вот этот вот комедию Мольера "Мещанин во дворянстве"? Что он был богат, разбогател, богаче многих дворян, и решил войти в высший свет. И там над ним постоянно смеялись, какие-то у него смешные ситуации происходили. То он там выпил воду, которую для того, чтобы пальчики смачивать перед тем, как салфетку взять, то там ещё что-то в таком, вилку не ту руку взял. В общем, постоянно он вызывал смех над собой. И он нанял учителей, и учителя его учили там и словесности, то есть как правильно говорить, и манерам поведения, и фехтованию, и, в общем, всему-всему. И вот учитель словесности, он ему объяснил, что есть проза, а есть стихи. И объяснил различия между ними. И этот господин Журден, он воскликнул: "Так что ж, выходит, что я всегда прозой что ли говорил?" Он говорит: "Да".

То есть, о чём, почему я это сейчас говорю об этом? Потому что мы многие вещи делаем, ну, мыслим в системе конструктов, короче говоря, но мы про это не знаем, что мы мыслим в системе конструктов. Понимаете?

Почему? Потому что мы не знаем, что такое конструкт. Значит, конструкт - это понятие, имеющее противоположные по смыслу полюса и спектр промежуточных понятий. А, так это мы знаем: близко-далеко, низко-высоко, узко-широко. То есть это пространственные конструкты, позволяющие нам ориентироваться в пространстве. Скоро или там не скоро это произойдёт, пространственно, это во времени.

Вот. То есть мы знаем такие конструкты, как пространственно-временные конструкты. Они позволяют нам ориентироваться в пространстве и времени. И, естественно, науки, физика, прежде всего, ну, математика, физика, они достигли больших успехов в использовании этих пространственно-временных конструктов в описании реальности. Ну началось всё с Декарта, который догадался, что геометрия, математика и арифметика, алгебра связаны. И, собственно говоря, об этом догадались в древности ещё. Вот, скажем, когда придумали линейки, стадии там на стадионах стали использовать. То есть мерить расстояние числами, то есть обозначать расстояние каким-то числом, сколько раз какой-то отрезок там укладывается в этом расстоянии. Ну то есть придумали линейки, рулетки и меры расстояния. То уже связали расстояние с числами, время с числами, когда календарь придумали.

Вот. Но Декарт сделал следующий шаг. Он понял, что можно числами обозначать положение точки на плоскости. Для этого нужно две линейки положить взаимно перпендикулярно вот так. И по одной линейке мы расположение этой точки перпендикуляр опускаем на одну линейку и на другую опускаем. У нас получается два числа. И вот эти два числа описывают положение точки на плоскости, а три числа - в пространстве, а четыре числа - в пространстве времени.

И вот вопрос возникает такой: а другие конструкты, какие мы знаем? Значит, я могу вам сказать, что почти все понятия наши являются либо конструктами, либо обозначают позиции на конструктах. Ну, например, цвет - конструкт спектр, да? Полюса конструкта - красный, фиолетовый. А промежуточное значение - вот эти цвета спектра. А конструктор вес - маленький, большой, возраст - молодой, старый там, бедный, богатый, красивый, уродливый там и тому подобное, умный, глупый, большой, маленький там и так далее. Ну, в общем, короче говоря, мы можем очень много вспомнить различных конструктов, которые мы в быту используем. Или, допустим, температура - низкая, высокая. Вот. Давление тоже низкое, высокое, освещённость - низкая, высокая. И какая-то есть шкала, в которой мы можем эти все величины измерять.

И то же самое касается и научных понятий. То есть почти все научные понятия тоже являются либо конструктами, либо позициями на конструктах. И вот мы, когда осмысливаем, да. Что такое полюса конструкта? Это какие-то состояния, которые сходны по смыслу и противоположны каким-то другим состояниям, относящимся к этому же конструкту, но в другом противоположном полюсе. Ну, допустим, температура холодная, горячая, да? То есть мы все, если будем классифицировать объекты, мы классифицируем их как холодные и горячие. И холодные будут на одном полюсе, горячие на другом. Тёплые там посередине, шкала Цельсия там или Фаренгейта, или Кельвина там между ними и так далее, да? А шкала может быть даже не числовой, а может быть порядковой просто, что больше, меньше там, ближе к холодному, тёплое там или ближе к холодному, прохладное там, вот такое. То есть это не количественная шкала, а порядковая, то есть там деградация текстовая.

6.3. Конструкты и мировоззрение

Значит, теперь, наше мировоззрение определяется тем, какие у нас конструкты. Их у разных людей разное количество. У некоторых людей есть такие конструкты, которые у других вообще отсутствуют. Если, допустим, при беседе используются какие-то понятия, которые у других людей отсутствуют, конструкты соответствующие, то они просто ничего не поймут. Значит, конструкты представляют собой оси. А дальше слушайте внимательно, оси в многомерном пространстве понятийном, в котором каждый объект внешнего, внутреннего мира имеет какое-то положение. То есть, допустим, берём какого-нибудь человека, к примеру, конкретного, у него определённый вес, возраст, пол, образование, национальность там и так далее, и так далее. И вот всё это, а все эти слова, вот национальность, там, пол, образование - это всё конструкты. Вес, возраст, всё это конструкты. Соответственно, это оси в пространстве. И на каждую ось можно опустить перпендикуляр или поднять с этой точки, соответствующей этому человеку. И они все пересекутся в некоторой области, соответствующей этому человеку. Этот человек будет изменяться, то есть у него рост меняется, вес меняется, возраст меняется, образование меняется, всё меняется. И получается некоторая, некоторая траектория в этом многомерном пространстве. Причём человек - это не точка в этом пространстве, а какой-то объём многомерный. Почему? Потому что все эти вот его параметры, такие как возраст и все прочие, они известны с некоторой точностью. Ведь возраст он известен, ну, обычно в годах он измеряется, там месяцы не пишутся. Поэтому сколько лет полных, да? Соответственно, у него по оси возраст будет год размер. А по оси вес у него будет размер, наверное, килограмм. Единица измерения килограмм, скорее всего, будет. И там тоже будет размер точностью до килограмма определён. И вот получается что? Что у него получится по каждой оси не точ, не линия и точка, а какой-то интервал и полосочка. И вот многомерный объём какой-то получится, который он занимает. А его эволюция будет трубочкой такой в этом многомерном пространстве. Вот он родился, имел маленький вес, маленький возраст, маленький размер, маленькое образование. Потом он рос, все эти параметры увеличивались, как-то всё это ползла эта кривулечка в этом пространстве. И вот она ползла, ползла и доползла до конца. Ну вы поняли, да? Вот. Так вот, также точно изменяются другие объекты. Тоже мы можем их изобразить в этом многомерном пространстве. Физики в теории относительности это называется интервал, то есть путь объекта в пространстве времени. Но объект при этом меняется, то есть не просто его положение меняется, он сам меняется.

И вот то, как мы понимаем предметную область, как мы, вот как исследователи, вы ж учёные, магистранты, будет будущие учёные, вы должны проводить исследования. То, как вы понимаете предметную область, определяется тем, какие у вас конструкты есть, позволяющие вам это делать. А вот когда вы пришли учиться, у вас одни были понятия, одни были конструкты. А когда вы стали учиться, у вас они изменились эти конструкты. Как они могут изменяться? Ну, во-первых, сама их система изменилась, то есть добавились новые конструкты. А во-вторых, система ваших понятий. То есть появились некоторые новые понятия, которых раньше не было. Они были сформированы на занятиях теоретических, как правило, ну и на практических закреплены. Вот. А те, которые уже были, ну, допустим, у вас был конструкт температура. Он был бытовой: холодно, тепло, там, горячо там, и от -60 до плюс там, там, не знаю, 30-40, например, там, вот такое что-то. А научный конструкт температура, он от нуля Кельвина до там триллионов градусов, которые внутри там нейтронных звёзд там или в точке термоядерного взрыва. Ну, в общем, короче говоря, смысл такой, что есть бытовой конструкт температура, а есть профессиональный. Они сильно отличаются по диапазону изменения величины. И бытовой может не быть шкалы числовой, а в профессиональном она может быть. То есть он точнее может измеряться. Вот. Соответственно, если мы возьмём в этом многомерном пространстве, которое представляет собой модель вашего мировоззрения, парадигмы реальности, то там все конструкты имеют диапазоны и какое-то их количество. То есть у вас можно сказать так, что ваше пространство мировоззрения вашего, оно имеет определённую размерность и определённый объём многомерный. И когда вы развиваетесь, знаете, что происходит? Когда формируются новые конструкты и диапазоны тех, кто имеется, увеличиваются. Что происходит с пространством вашего мировоззрения, с когнитивным пространством? Кто-нибудь скажите, как вы думаете?

Ну давайте, смелее говорите, высказывайте гипотезы.

Что, никто не может предположить, что происходит с пространством? Размерность его увеличивается, и объём его увеличивается.

6.4. Размерность пространства и проекции

И, значит, я могу вам сказать, что такое размерность пространства. Вот представьте себе, что у вас есть трёхмерная модель какого-то реальности. И вы её фотографируете. То есть осуществляется проекция трёхмерного этого объекта на двумерное пространство, плоскость. Вот, фотография, она, может быть, даже очень хорошая, но она, так скажем так, фотография необратимо теряется информация о третьей координате, потому что она сворачивается в ноль. Проектирование, проекция происходит трёхмерного в двумерное. Вот представьте себе, два человека беседуют, у одного там размерность пространства когнитивного там 375.000, а с другим он беседует, а у него размерность пространства всё 10, там или 70. То есть в 1000 раз отличается размерность. То есть если, допустим, этот человек, у которого размерность высокая, что-то скажет, то для того, чтобы тот понял, ему нужно это всё спроектировать в, ну, грубо говоря, в плоскость, понимаете, так упростить. Вот. И тогда тот въедет, у кого размерность меньше.

Вот и все дела. То есть конструкт - это очень важное понятие для, то есть это основа, ось пространства многомерного, которое, в котором, значит, существуют образы всех наших объектов внешне, внутренне реальности. Где-то в разных областях этого пространства они там существуют. Оно позволяет нам ориентироваться и во внешнем мире, и во внутреннем мире.

7. Применение к анализу факторов в садоводстве

Так вот теперь смотрите, ребята, вот у нас все значения факторов, которые есть в модели, были проанализированы, как они влияют на объект моделирования. И определено, какие факторы, значения факторов сходно влияют, какие сильно отличаются.

Галлюцинация. Ну вот. То есть те факторы, которые оказывают сходное влияние на объект моделирования, объединены в кластеры на низком уровне различия. Потом дальше, дальше, и потом две группы значений факторов, вот эта синенькая - это обуславливающие переход на один полюс конструкта объекта моделирования, а эти - на другой полюс конструкта. То есть эти факторы оказывают противоположное влияние.

Чем это важно? Что мы знаем, какие факторы оказывают сходное влияние, а какие сильно отличающиеся влияние. Чем это важно? Это важно потому, что когда мы получили, ну это график межкластерных расстояний классический. Здесь кластерная формула, которую я придумал, как писать в такой форме, уровень вложенности. Сначала эти объединяются, эти, потом они друг с другом и с этим объединяются и так далее. Потом вот эти объединились, к ним потом добавляется этот, потом ещё вот этот добавляется. То есть вот вы поняли. То есть здесь очень наглядно изображена вложенность кластеров этих друг в друга, матрёшечный стиль, так сказать, это.

8. Система поддержки принятия решений "Эйдос" (Развитый алгоритм)

8.1. Преимущества перед SWOT-анализом

И вот что я могу сказать, что всё это очень важно для принятия решений. Мы с вами рассмотрели упрощённый вариант принятия решений 448. Что в нём упрощённого, ребята? Во-первых, только один класс. Мы выбрали один целевой класс. Во-вторых, мы посмотрели, какие значения факторов рекомендуются использовать, а какие не рекомендуются, категорически не рекомендуются использовать. Посмотрели на них, на эти факторы. Допустим, здесь есть факторы агротехнологии, ребята, или факторы окружающей среды. Они все влияют существенно на результаты выращивания. И вот мы смотрим, а какой-то фактор агротехнологии мы не можем использовать по финансовым соображениям. То есть, чтобы его использовать, нам нужно закупать какое-то средство защиты, средство его внесения, хранилище строить специальное для его хранения, строить станцию обслуживания этой техники, которая будет носить это средство защиты с помощью GPS-навигации, направлять в Германию специалистов для того, чтобы со знанием статьи языка хорошим. То есть продвинутый бытовой уровень. То есть они должны без акцента говорить бегло по-немецки. Вот, на бытовые темы. И вот, значит, их туда посылают, они получают там подготовку, сертифицируются. И после этого они могут обслуживать эту технику, ремонтировать её. Вот. И, в общем, всё это, если посчитать, то думаешь, да ну его, лучше не связываться с этим.

8.2. Учет ограничений и замена факторов

Возникает вопрос: а если мы этот значение фактора уберём, вот оно нам рекомендует его использовать, а мы его возьмём и скажем: "Нетушки, мы не будем его использовать". Вот это значение фактора мы не будем использовать. А мы, причины у нас есть свои, то есть по которым мы так решили. Возникает вопрос: а тогда получится у нас или нет достичь целевого состояния?

Значит, ребята, значит, идёт речь о том, что нужно принимать решения в системе, в которой мы ставим цели, то есть нам ставят цели руководство, и мы должны эти цели достигать. А объектом управления у нас является биосистема. Ну, как говорят, искусственная экосистема. Ну, допустим, сад. Вот. Искусственный лес, так, грубо говоря, из фруктовых деревьев. И вот возникает вопрос: как это сделать?

Значит, я хочу вам сказать сейчас очень важную вещь. А вы сейчас слушайте и потом мне ответите, услышали, что я сказал, или нет. Значит, по сути дела, мы все своего рода господина Журдены. Вот в частности, с ну, тот факт, он говорит о том, что вы должны добиваться достижения определённых целей. Значит, я могу вам сообщить, ребята, такую приятную новость. Сообщаю вам, что я кандидат технических наук по специальности 051306 - автоматизированные системы управления и доктор экономических наук по специальности математические инструментальные методы экономики, что одно и то же практически, по которой присваиваются технических, физмат или экономических наук. То есть я специалист по теории управления, короче говоря. И у меня, как вы слышали, наверное, я вам, по-моему, говорил про это, в России первый рейтинг в этой области, в области автоматизированных систем управления и систем искусственного интеллекта.

Вот, видите, кибернетика. К ней относятся и системы искусственного интеллекта, и системы управления. Первый рейтинг в России. Ясно, да, кто вам преподаёт? Чтобы вы так себе представляли чуть-чуть.

Именно вот в этой области. Так вот, применение интеллектуальных систем для управления. Понимаете? Вот. И вот я вам могу сообщить, что вам, когда преподают ваши дисциплины, то вам никто не говорит, что вы специалисты по теории управления, и объектом управления у вас является искусственная экосистема, а в качестве управляющих факторов вы используете различные агротехнологии, а также сам сорт. Вот, определяет свойства его, определяют результаты. И влияние факторов окружающей среды на сад существует, и вы должны выбрать такие микрозоны выращивания, где факторы окружающей среды действуют благоприятно. Иначе вам ничего не удастся сделать с помощью ваших управляющих факторов, если вы там в тундре будете выращивать абрикосы. Вот, ну и так далее. Значит, и вот здесь вот всё это описано, понимаете? Здесь описан цикл управления в замкнутой системе управления, в которой свойства объекта управления определяются его предысторией и его текущим состоянием. Информация об этом поступает в управляющую систему. В ней осуществляется синтез и верификация модели, выработка управляющего воздействия путём однократного решения обратной задачи прогнозирования с SWOT-анализ, который я вам показал. Вот.

И несколько раз также решается задача прогнозирования и используются результаты кластерно-конструктивного анализа. Как это делается, ребята? Вот развитый алгоритм принятия решений в системах управления на основе СК-анализа системы Eidos, интеллектуальных системах управления. Сначала ставятся цели управления, определяется одно или несколько целевых состояний объекта управления. Я хочу подчеркнуть, в SWOT-анализе только одно целевое состояние обеспечивается. То есть мы не можем здесь сразу два задать. Вот SWOT-анализ - это метод экономический управления, который как раз и представляет собой упрощённую форму принятия решений, где есть факторы, которые рекомендуется использовать, есть факторы, которые категорически запрещается использовать, потому что они препятствуют достижению цели управления. SWOT-анализ имеет несколько ограничений. Первое ограничение, что только один класс может быть задан в качестве целевого состояния. Второе, значит, здесь перечислены факторы, а вот эти вот нагрузки, сила влияния этих факторов и знак влияния определяются путём экспертных оценок. Я ни разу не видел защиты ни магистерской, ни там дипломной, ни магистерской работы, ни кандидатской, ни докторской, где бы вот эти весовые коэффициенты расчётным путём определялись. Всегда на основе экспертных оценок. Причём обычно экспертами выступают сами те, кто писал эту работу. И они далеко не всегда являются экспертами. А эксперты определяют вот эти весовые коэффициенты на основе интуиции, опыта и профессиональной компетенции, то есть вообще от фонаря, грубо говоря. Так вот им кажется, так они и определяют. А кажется им не всегда правильно, потому что одно дело действовать, а другое дело говорить об этом действии. Не все эксперты имеют способность, развитую способность вербализации. Есть прекрасные практики, которые не могут объяснить, почему они что они делают и почему они это делают, и как получается, что таким они достигают результата. Вот, скажем, Мичурин, он не написал таких вот фундаментальных научных работ, где бы объяснялось, как он это всё делал, понимаете? То же самое Никола Тесла.

Вот. То есть есть практики, вот я помню, был академик Мальцев такой, народный, который был, ну, непревзойдённым он был хлеборобом, то есть вообще удивительным, великолепно он это, или не хлеборобом, а в какой-то другой области сельского хозяйства он работал. Вот. Но он не был теоретиком, он не мог объяснить, почему он так делает. Он всё это чувствовал интуитивно, понимал хорошо, вот, но не мог описать и обосновать.

Вот. Так вот, что получается? Что мы должны обобщить этот SWOT-анализ на случай, когда у нас несколько классов указано. И и объяснить, что нам делать в случае, если какие-то факторы мы не можем использовать по тем или иным причинам. Ну, обычно либо финансовые, либо технологические. Они тесно взаимосвязаны. Если финансы не ограничены, тогда можно купить любую технологию и применить. Но тогда это может быть убыточно, понимаете? Поэтому, возможно, в этом нет смысла.

8.3. Алгоритм действий в Eidos

Так вот, если мы ставим цели управления, определяем одно или несколько целевых состояний объекта управления. Значит, эти целевые состояния обычно выражаются в натуральном выражении - это обычно количество и качество продукции, а в стоимостном выражении - это обычно прибыль и рентабельность производства этой продукции, продажи. Потом проводим то, что вот я уже показывал, когнитивно-целевую структуризацию предметной области. То есть мы определяем, что мы в качестве факторов будем рассматривать, а что в качестве результатов влияния этих факторов. Потом проводим формализацию предметной области, то есть мы разрабатываем справочники прошлых и будущих состояний и с их помощью кодируем исходные данные, получаем обучающую выборку или тренировочную выборку. Потом определяем, создаём, осуществляем синтез и верификацию статистических и системно-когнитивных моделей, определяем наиболее достоверную из них по F-критерию Ван-Ризбергена и его обобщениям. Я предложил нечёткий, мультиклассовый обобщение критерия Ван-Ризбергена, инвариантно относительно объёма выборки. И потом смотрим, у нас одно задано целевое состояние или несколько? Если одно, то да, здесь дальше нам делать нечего, мы переходим на следующие там этапы. Если же несколько, то слушайте внимательно. На шаге четыре мы оцениваем корректность поставленных целей путём сравнения системы детерминации целевых состояний методом когнитивной кластеризации или просто на основе матрицы сходства. То есть определяем, являются ли целевые состояния совместимыми, то есть достижимыми одновременно по обуславливающим их значениям факторов, или они являются взаимоисключающими, альтернативными.

Здравствуйте, Людмила Владимировна, я занятие веду.

Евгений Вениаминович, как вы там поживаете?

Да ничего так, более-менее. Занятие веду, у меня пять пар подряд сегодня.

Да, вы не едете в Болгарию? Я хотела спросить вас вопрос. Вот скажите, примерно когда статья, ну, появится, имеет смысл сейчас её включать в эту систему, смотреть?

Конечно, конечно. Она именно за третий квартал, в сентябрьском номере. А вы мне пришлите письмо, я вам пришлю подтверждение, вы туда его вставите в отчёт, обоснование и всё.

Спасибо огромное.

Ага, давайте.

Профессор Цаценко. Наверное, вы знаете её. Да, ребята?

Вот когда вы не отвечаете на мои вопросы, то у меня возникает...

Нет, я нет. Я нет, честно.

Не знаете Цаценко?

Нет. А что?

Людмилу Владимировну?

Нет.

Ну ладно. Ну, ну, нет, так нет.

Ну я первый год учусь, просто я не знаю.

Ну, понятно. Значит, так вот эти состояния будущие, они являются либо достижимыми одновременно, либо взаимоисключающими, альтернативными по системе детерминации или по по системе факторов, которые обуславливают переход в эти состояния. И вот возникает вопрос: а они могут быть достигнуты одновременно эти состояния или нет? Если не могут, ребята, тогда мы возвращаемся на самое начало и говорим о том, что цели управления надо ставить более корректно. Ну нельзя одновременно, одновременно, так сказать, двигаться и на запад, и на юг, понимаете? То есть, вернее, на запад и на восток. Ну, вернее так, вот Колумб доказал, что это возможно, но очень большой объезд. Очень крюк большой.

Вот. Ну, не Колумб это доказал, а это доказал Магеллан. А Колумб думал, что это он докажет, но не удалось. Так вот, он только открыл другой материк, а не вокруг Земли проплыл. Так вот, ребята, значит, если состояния одновременно достижимы, тогда можно решать задачи одновременного достижения этих состояний. Для этого мы решаем задачу поддержки принятия решений в упрощённом варианте путём автоматизированного когнитивного SWOT-анализа. Так вот система Eidos, я вам почему говорил, что там эти коэффициенты экспертами определяются? Почему? Потому что система Eidos является единственной системой на данный момент в мире, которая автоматизирует SWOT-анализ. Я вам приводил информацию об этом?

Сейчас, секундочку. Значит, я сейчас попробую вам дать ссылочку на статью про SWOT-анализ. Я специально поиск осуществил, пытался найти такую систему и ничего не нашёл. Единственное, нашёл систему Парус, в которой мы можем ввести вручную значения факторов, весовые коэффициенты вот эти, вручную всё ввести. Она тогда нарисует вам SWOT-диаграмму, SWOT-матрицу. Ну это же не то же самое, что разработать эту SWOT-матрицу. А если вы не знаете, что туда вносить? Ну вот система Eidos, она рассчитывает SWOT-матрицу на основе модели, понимаете, созданной на основе эмпирических данных.

Потом дальше мы оцениваем технологические и финансовые возможности применения на практике рекомендованных на шаге значений факторов. То есть для всех этих состояний проводим этот SWOT-анализ, и все эти значения факторов выписываем. И смотрим, можем ли мы на практике применить эти значения факторов? Если можем, то вопросов у нас вообще нету матросов, тогда мы вот перемещаемся на самый выход из модели. То есть нам тогда из процесса принятия решения, тогда нам делать больше здесь нечего. То есть осталось только применять это на практике эти значения факторов.

А если у нас нет возможности применить всю эту систему факторов, тогда мы исключаем из системы значений факторов, которые рекомендованы были на шаге шесть, те из них, которые у нас нет возможности использовать по тем или иным причинам. И потом мы что делаем, ребята? Прогнозируем, прогнозируем результаты применения на практике сокращённой системы значений факторов. А мы с вами уже проходили идентификацию и прогнозирование. Правда, я вам не рассказал, как принимается решение о принадлежности к классу. Сейчас я вам расскажу об этом. И вот, сейчас я вспомнил про это, что не рассказал. И тогда мы смотрим, а если мы этот фактор не будем или несколько факторов использовать, достигнем мы целевого состояния или нет? Если достигнем, тогда это нас устраивает, тогда мы выходим на выход из алгоритма принятия решения. А если мы не можем достичь, то есть если мы удаляем какие-то факторы, которые мы не можем использовать из рекомендованных SWOT-анализом, и система тогда не достигает целевого состояния, тогда возникает такая мысль, что если этого не происходит достижения целевого состояния, тогда нужно какие-то эти вот факторы, которые мы не можем использовать, заменить какими-то другими. Вот. И мы можем их заменить факторами, а дальше слушайте внимательно, которые оказывают сходное влияние на объект управления, но такими, которые есть возможность использовать. Ну, например, вам дали задание вырастить на каком-то поле большой урожай пшеницы. А вы смотрите, на этом поле был подсолнух в прошлом году. А вы знаете закон севооборота, что там после него вообще ничего не растёт. То есть, причём 10 лет. То есть, что делать? Вы начальнику своему, хозяину хотел сказать, руководителю говорите: "Это невозможно. Вы какое-нибудь другое поле выберите. На этом поле надо ему восстанавливаться, ему ещё несколько лет надо под парам быть, пока оно восстановится. Причём надо его удобрять там и так далее". Он говорит: "Ну если ты ещё раз скажешь мне, что ты что-то не можешь, я тебя уволю, найму другого, который этого мне не будет говорить. Вот давай, достигай этой цели, я тебе поставил цель, работай. Дальше твоё дело, ты специалист в этом". Я говорю: "Так я потому как специалист и говорю, что это невозможно". Вот ещё раз скажешь, уволю. Всё, иди. Вот. И, короче говоря, сейчас руководство обычно не спорит, когда ему что-то там предлагают или там что-то говорят, они сами знают всё. Уходит этот агроном с поникшей головой. И пока идёт, он вспоминает вдруг, что когда он был магистрантом, то профессор Луценко рассказывал, что оказывается, ж можно ж какие-то факторы заменять другими, сходными по влиянию на объект моделирования. А для этого надо узнать, какие же факторы оказывают сходное влияние? Для этого нужно провести э кластерно-конструктивный анализ факторов. Вот эту дендрограмму факторов посмотреть, которую я вам только что показывал. Вот. И выбрать такие факторы, которые в одних кластерах находятся. 4 3 2 3. Либо э даже не на дендрограмму посмотреть, а просто на матрицу сходства и различия факторов. И найти вот этот фактор, который у нас нет возможности использовать, и какие факторы сходны по его по влиянию на объект моделирования. Какие высокий уровень сходства фактора с ним имеют? И попробовать их применить. И опять же, то есть заменить этот фактор вот этим, например, этот вот этим. Ну здесь факторы у нас не агротехнологические, здесь факторы морфологические свойства, их не заменишь. А вот э ну, в некоторых случаях, может быть, и можно. А технологические можно. Вот если, допустим, мы узнали, что там э не было бобовых предшественников, а были другие. Возникает вопрос: а какие факторы оказывают сходное влияние примерно такое же на объект моделирования, как бобовые предшественники?

Ну, насколько я знаю, это внесение азотных удобрений. Почему? Потому что бобовые предшественники, они как раз вносят азот в почву естественным образом, да? Вот. То есть если мы внесём азот в почву, то есть надежда, что мы получим результаты такие, как будто там были бобовые предшественники. Вот о чём речь идёт. Но при этом есть некоторые детали. Значит, нужно иметь в виду...

Значит, заменяем, заменяем бобовые предшественники внесением удобрений и смотрим, получится результат или нет. И у нас получается положительный результат. Тогда, если сформированная система факторов путём замены тех, которые мы не можем использовать, и совмещения нескольких систем факторов с состояниями будущими, которые сходны по этим своим факторам, что можно поэтому их совместить, обеспечивает достижение целевого, тогда выход из алгоритма принятия решения. Если же нет, то опять идём на начало. И вот здесь не только цели управления можем поменять, но и меняем саму модель. То есть используем другие факторы, более широкий набор факторов используем в модели, которые нам обеспечат достижение целевого состояния, может быть. То есть мы понимаем, что в этой системе факторов это не получается сделать. Вот. Но что касается удобрений. Значит, здесь, если вы посмотрите на дендрограмму, которую я показывал, ну только для этой задачи, которую я рассказываю, то тогда вы увидите, что там не только внесение удобрений, но и определённые способы вспашки. Если специалиста спросить, что эти такие способы вспашки обеспечивают? Почему там вот их влияние сходно с бобовыми предшественниками, внесением азота в почву? Именно вспашки. Оказывается, что просто так, если внести эти удобрения, то ещё неизвестно, что с ними там станет. После первого дождя они там ручейком стекут куда-нибудь в речку, и на этом всё закончится, и рыбы подохнут. Вот. А если, значит, спахать определённым способом, и борозда будет направлена определённым образом по склону местности, и там вспашка будет там без разрушения какого-то слоя, там дисковая. Ну сейчас немножко фантазирую. Вот. Я этого не знаю. В общем, там есть, я сейчас не помню просто. Там я решал эту задачу, но там какие-то были определённые способы вспашки, которые, как мне сказали специалисты, они препятствуют вымыванию или диффузии удобрений из почвы. То есть их не вымывает, понимаете, во время дождей. Они там остаются. Вот. Потому что если просто внести, они исчезнут и всё через какое-то время, через неделю там. Эффект их действия будет слаб. Если же, значит, этот способ вспашки одновременно с внесением таких удобрений применяется, то тогда они действуют эффективно, примерно как вот предшественники.

Вот примерно такие вот дела, ребята. То есть мы можем принимать решения не просто вот на основе SWOT-анализа, а развитый алгоритм принятия решений вот здесь применять вот в этой системе.

8.4. Адаптивность системы

Предыстория - это как раз и есть вот севооборот, в принципе. Если объект управления - это искусственная экосистема, там выращивание каких-то культур. Факторы окружающей среды - это микрозона выращивания, почвенно-климатические факторы, освещённость, увлажнённость, заморозки там на различных фазах развития, когда там цветение, там заморозков не должно быть, как правило. Иначе будет рискованная зона выращивания.

Ну, понятно. Вы, наверное, знаете лучше меня. То есть я сейчас просто сказал, ребята, что вы и мы тоже, как господин Журден, не знаем, оказывается, что мы, оказывается, специалисты по теории управления. И мы решаем задачи достижения целей, выработки управляющих воздействий. И делаем это с использованием интеллектуальных систем, которые применяются в этих интеллектуальных системах управления. А почему эта система адаптивная, ребята? Потому что можно старые данные удалять из базы исходных данных, на основе которой формируется модель, а новые добавлять. И в результате у нас модель будет отслеживать динамику предметной области.

Вы, наверное, слышали, что у нас сейчас потепление климата происходит, и все эти вот климатические пояса ползут вверх, смещаются. И где-то уже лет пять мне кто-то из специалистов с плодфака говорил, что исследовали факторы климатические на Кубани. И где-то лет пять назад по критериям мы перешли в зону субтропиков, влажных субтропиков. То есть получается, что вот те, та зона, которая была в Сочи, не в Сочи находился в этой зоне, Абхазия, она переползла сюда. Значит, я помню, в юности я гулял по Геленджику, и мы видели там деревья, эти, где с розовыми цветами, которые нас восхищали, это, как его называют, мимоза. Вот. А у нас здесь таких деревьев не росло. А сейчас вот во дворе у нас они растут. Я к чему это клоню? К тому, что у нас зоны сместились, и зоны, деревья, которые росли в южной зоне, они оказались у нас. И у нас они прекрасно себя чувствуют, и цветут, и плодоносят, и всё. А раньше это было невозможно. Соответственно, актуальной становится задача разработки сортов, которые бы росли бы в наших условиях эффективно, возделывались и так далее. То есть с учётом этого, динамики окружающей среды. И вот такие системы, которые я сейчас рассказываю, они способны учитывать эти изменения, потому что они строят модели на основе фактических результатов воздействия на объект управления и фактической реакции объекта управления. То есть они учитывают и какая климатическая среда, почвенно-климатические условия, и освещённость, и температурный режим, увлажнённость, и учитывают, какие факторы мы оказываем управляющие на эту, на этот объект управления, и что в результате получаем. И всё это учитывается в модели. И в следующий раз она оказывает более адекватное воздействие на объект управления с учётом, ну, то, которое более эффективно переводит его в целевое состояние. То есть система может работать в режиме обучения, то есть она сама, то есть модель управления сама обучается.

8.5. Роль интеллектуальных систем

Вот. И это называется всё автоматизированная система управления. То есть решения принимаются с участием человека, но с помощью автоматизированных интеллектуальных систем. То есть вы можете это использовать. То есть не просто посчитать, что что-то на что-то как-то влияет, а использовать для развитой формы управления в адаптивном режиме. Это совершенно меняет уровень. То есть чаще всего те, кто защищаются, они даже не представляют себе, что они занимаются управлением. Что есть управляющие факторы, есть объект управления, модели должны быть, которые обеспечивают принятие решений и так далее, и так далее. Я вам скажу парадоксальную ситуацию такую существует. Я вот член совета по экономике, где есть специальность 082005 - управление в народном хозяйстве по отраслям. И управление предполагает несколько вещей: что есть объект управления, управляющие факторы, информация обратной связи, есть управляющая система, где принимается решение об управляющих воздействиях, есть цели управления, есть факторы окружающей среды, влияющие на объект управления. Я спрашиваю, ну, я не спрашиваю, потому что я не вредный человек. Если бы я это спросил, я бы завалил бы этого защищавшегося там соискателя степени, чего я делать не хочу и не собираюсь. Но я сижу и вижу, что человек вообще ничего не сказал об управлении. У него специальность 082005. Я хотел бы его спросить: "Вот это управление?" Он говорит: "Да". Я говорю: "Хорошо. Теория управления предполагает, что есть объект управления. Что у вас является объектом управления?" Ну, тут он может ещё что-нибудь скажет там, что там хозяйство какое-нибудь там или район, или подотрасль какая-то АПК. А я ему тогда спрошу: "А управляющие факторы какие у вас? Вот вы свою работу провели, вы должны нам рекомендовать, какие управляющие факторы использовать для воздействия на объект управления". И всё, и он, значит, то есть они таких слов вообще не знают, они их не произносят. А если он скажет, какие управляющие факторы, то тогда скажу: "А как вы определили, что эти управляющие факторы будут оказывать именно то воздействие на объект управления, которое переведут его в целевое состояние? Как вы это решили? С какого перепуга?" Это можно решить только на основе обобщения опыта. И должна быть модель математическая, которая этот опыт обобщает. И должна быть программная система, которая обеспечивает работу этой модели, её использование. Как вы это решили, что у вас вот эти решения ваши повысят там чего-то там на какой-то процент? Количество, обычно говорят, количество продукции повысится. С какого перепуга оно повысится? А вы это проверяли? А ну, они даже не знают, что решение принимается на основе модели. Ну если, допустим, даже они это знают, я их могу спросить: "А какова достоверность этой модели? Вы её проверяли?" А вдруг эта модель недостоверна, а вы на основе неё вырабатываете рекомендации?

То есть, понимаете, возникает масса вопросов вот таких вот. А как принималось решение у вас, соответственно с моделью? И вот всё, понимаете, они ничего этого не знают вообще. И то же самое касается, наверное, плодоводства и садоводства. То есть люди занимаются управлением сложнейшими экосистемами, которые они сами же и создают. И они совершенно об этом не подозревают, что они занимаются управлением. Что там есть объект управления, управляющие факторы, модели должны быть, которые обеспечивают принятие решений и так далее, и так далее. Сейчас я вам продемонстрировал, ребята, что система Eidos является системой, которая полностью обеспечивает решение всех задач, которые необходимо решать в управляющей системе, в случае, когда у нас применяется развитый алгоритм принятия решений при переводе объекта управления в заданное целевое состояние. Вот что я хочу вам сказать. Вы, наверное, это уловили.

9. Заключение и перспективы

Теперь давайте посмотрим, как решается задача идентификации и прогнозирования. Я вам начал рассказывать, но не всё рассказал. Значит, обратите внимание, здесь справа есть два окошка: верхнее и нижнее. В верхнем окошке семантический резонанс знаний написано, интегральный критерий, а в нижнем - критерий сходства, сумма знаний. Кликаем по кнопочке. А обычно никто не догадывается, что это кнопочка, но это кнопочка. Интегральный критерий сходства, сумма знаний.

Значит, теперь слушайте внимательно. Значит, мы в модели, которые у нас созданы, ну, допустим, модели, смотрим в режиме 5.5, в модели F1, мы видим, что в битах у нас есть информация, то есть знания, или, скажем так, у нас в модели отражено, какое количество информации содержится в каждом значении фактора о том, что объект, на который действует это значение фактора, объект управления, перейдёт в определённое состояние будущее, соответствующее классу. Прямо в битах это написано. А теперь представьте себе, что на объект моделирования и управления действует не одно значение фактора, а несколько, много значений факторов. Тогда как это понимать? В какое состояние перейдёт объект моделирования?

И есть лемма Неймана-Пирсона, чрезвычайно убедительно звучащая, что, они правда её доказывали для хи-квадрат критерия, но он тесно связан, как я вам показывал, он тесно связан с критерием информации Харкевича. То есть фактически хи-квадрат - это то же самое, только это путём вычитания определена сходство фактической и теоретической частоты наблюдения признаков, а в мере Харкевича путём деления это определено. И всё, понимаете, разницы, в основном разницы-то нет, это очень сходные по смыслу критерии. И они, значит, сформулировали э Нейман и Пирсон. Нейман вообще гениальный математик, Пирсон тоже, потому что он э практически, ну, ну, скажем так, внёс основной вклад в развитие параметрической статистики, или можно даже сказать, что он её единолично разработал параметрическую статистику.

Вот. Значит, теперь слушайте интересный момент. Они посчитали, что, доказали, вернее, такое следствие из теоремы, из одной из теорем статистики, что объект относится к тому классу, о принадлежности к которому в его системе признаков больше всего информации. Понимаете? То есть получается так, что если мы о каждом признаке знаем, какое количество информации в нём о принадлежности, не принадлежности к тому или иному классу, а у нас известно, что признаков-то столько-то вот у объекта, и такие-то вот они эти признаки, то мы можем сказать, что объект перейдёт в то будущее состояние, о переходе в которое в его системе признаков больше всего информации. Как это можно посчитать это количество информации? Для этого нужно знать, какие признаки есть, какие отсутствуют у объекта, который мы идентифицируем. Значит, чтобы это знать, мы, значит, формируем массив, я его назвал массив L, локатор, который э имеет элементов столько, сколько вообще есть признаков в модели. То есть все градации всех описательных шкал, э все там подряд идут. Вот. И если этот этот признак есть у объекта, то в этот массив элемент этот соответствующий имеет единицу, то есть равен единице. А если этого признака нет, то нулю равен соответствующий элемент. И просто, значит, мы можем взять теперь циклы по всем признакам и смотреть, есть он или нет. Если он есть, суммируем количество информации о принадлежности к житому классу, к определённому классу. А G меняется от одного до конца, сколько всего там классов. То есть мы для каждого класса это делаем. Ну, берём какой-то класс и определяем, какое количество информации о принадлежности к нему в признаках данного объекта, суммарное. Если признак есть, проверяем, единичка или нолик. Если единичка, то суммируем, если нолик, не суммируем. А можно не проверять ничего, ребята. То есть не писать этой команды if then else там, всё это вот в программе, а можно просто взять и перемножить два вектора: вектор класса и вектор объекта, состояния объекта. Что это за вектора? Я ещё раз вернусь к модели. Лучше несколько раз показать, чтобы было вполне понятно всё. Вот смотрите, вот модель. Колоночка описывает класс. Вот эти вот числа - это координаты этого вектора. Значит, у них есть название этих векторов, координат, извините, и значение. Вот. Так вот, есть вектор класса таким образом, есть вектор объекта, и они просто перемножаются друг с другом, вычисляется скалярное произведение. То есть два вектора умножаются друг на друга скалярно.

Вот. В результате у нас получается суммарное количество информации в признаках какого-то класса о принадлежности к житому классу, не, в признаках какого-то объекта о его принадлежности к житому классу. Это вот и есть скалярное произведение в координатной форме. Это интегральный критерий, который применяется в системе Eidos, который я назвал сумма знаний. Это интегральный критерий, потому что у нас в матрицах, системно-когнитивных матрицах, у нас не информация содержится, а знания, позволяющие нам принимать управляющие решения, достигать цели, то есть по определению это знания. А соотношение смысловом данных, информации и знания я вам, наверное, рассказывал всё-таки, да? О концепции смысла Шенка и Абельсона. Я вас уже спрашивал, рассказывал я или нет? И вроде вы сказали, что рассказывал. Ну как-то так нечётко это прозвучало, я не совсем понял. Значит, рассказывал или нет? Рассказывал, нет, ребята?

Нет, вроде не рассказывали.

Вроде нет, вроде бы так вспоминаю.

Я рассказывал соотношение данных, информации и знания, да?

Нет, нет.

Понятно, понятно. Ну сейчас тогда расскажу очень коротко. Время вроде как есть. Так вот этот критерий, чем он интересен? Тем, что он не зависит от того, являются ли вектора классов взаимно перпендикулярными. Если являются, тогда пространство является ортонормированным, и можно применять расстояние Эвклида, ну, теорему Пифагора, только многомерный вариант. Корень квадратный из суммы квадратов разности координат этих двух векторов для измерения расстояния между ними. Но если пространство не ортонормировано, а у нас оно именно такое, и оно всегда такое на самом деле. Ортонормированное, то есть когда взаимно перпендикулярные оси пространства - это просто абстракция. Ну такая же, как вот материальная точка там и математическая точка. То есть это реально в природе не встречается. И мы видели, что они все взаимосвязаны эти классы и все факторы. И таким, раз факторы взаимосвязаны, тогда получается из этого вывод, что система нелинейна. И, в общем, это означает, что параметрическая статистика не подходит для описания этой системы, а только ранговая, непараметрические методы, теория информации подходит. Понимаете? То есть мы сейчас вот я вам рассказываю, что у вас дисциплина называется матмоделирование, да? Вот. И обычно понимают так, что матмоделирование, анализ данных - это понимают обычно под этим статистику. И обычно понимают под этим параметрическую статистику. А я вам рассказываю, что она неприменима в вашей области. Она применима в области механики, где вот мы можем взять какое-нибудь тело и действовать на него двумя силами. И по закону Ньютона у него будет ускорение там, М на А. Вот. И при этом, если взять силы две, которые под каким-то углом действуют на это тело, то тогда они складываются по методу параллелограмма. То есть их действие является аддитивным на объект управления, и суммарное действие двух сил является суммой их действия по отдельности. Это определение линейной системы. У вас же так не получается в вашей области. То есть системы, которые связаны с растениями, с животными, с людьми, они являются существенно нелинейными. Вот. Так вот, методы параметрической статистики вот такие вот влияния факторов не обрабатывают, только вот в начале около нуля, где почти линейно меняется, вот это они обрабатывают, вот этот кусочек вот. А когда сильно влияют факторы, они не, то есть это всё слабое влияние факторов. Когда слабо они влияют, тогда параметрическая статистика пригодна. Когда сильно, это уже непригодным оно становится. Значит, ещё хочу вам сказать, что то же самое касается и физики. Если, допустим, два пучка света, два лазера, вот так вот лучи пересекаются, есть точка, где они соприкасаются, то потом они расходятся так, как будто бы второго луча не было. Но это до тех пор, пока не очень высокая интенсивность этих лучей. Когда она становится очень высокой интенсивность, высок, очень высокая мощность лазеров, то лучи света начинают между собой взаимодействовать. То есть это становится нелинейная система возникает, и она уже не описывается классическими вот методами электродинамики, то есть там надо уже что-то такое выдумывать там, как это всё э рассчитать и спрогнозировать. Так вот система Eidos, она способна обрабатывать, создавать модели нелинейных объектов. То есть в ней это корректно. Я вам показывал там примеры я приводил, когда когда несколько факторов действуют на систему, и результаты их совместного действия не являются суммой их результатов действия по отдельности, а влияние каждого фактора уменьшается. И поэтому, когда мы анализируем с помощью системы Eidos влияние нескольких факторов, то у нас получаются корректные результаты, в отличие от методов параметрической статистики. Поэтому я вам и говорю, что она практически неприменима. То есть если вы хотите применить параметрическую статистику, вы должны доказать, что у вас гаусовское распределение там наблюдается. И факторы не зависят друг от друга. Вот. Ну, и будущие состояния друг от друга не зависят. Этого вы не докажете, потому что они зависят друг от друга. А если вы этого не будете доказывать вообще, тогда значит, вы вообще некорректно применяете параметрические методы. А если вы это доказываете, и потом у вас получается, что это не удалось доказать, то тогда надо применять другие методы. А какие другие методы? Вот другие методы - это система Eidos. И методы непараметрической статистики можно применять ранговые. Но я вам рассказывал про развитие статистики и говорил, что развитием непараметрической статистики являются интеллектуальные технологии.

Так вот, вот это вот, этот интегральный критерий, скалярное произведение, он является корректным в случае, если пространство является неортонормированным, каким оно всегда является, то есть факторы взаимодействуют, и классы взаимодействуют. И нелинейным таким образом является объект моделирования. Оно является, это мера является корректной. Это интегральный критерий. Кроме того, это определение белого шума. То есть шум - это такой сигнал, корреляция любого фрагмента которого с любым другим фрагментом которого стремится к нулю при увеличении размеров этих фрагментов. То есть если фрагмент увеличивается, то корреляция стремится к нулю. То есть это что означает? Что один элемент этого сигнала не зависит от другого элемента этого сигнала. Понятно, да? Вот. Так вот, если у нас в исходных данных есть шум, то тогда вот этот вот интегральный критерий, его подавит этот шум, потому что это и есть определение белого шума. То есть получится, что шум в среднем равен нулю, а сигнал вытащится оттуда. Вот. То же самое касается мер знаний. Сам хи-квадрат тоже для этого предназначен. Количество информации будет равно нулю. Если мы возьмём, посчитаем вот в этих моделях, вот, которые здесь у нас есть, посчитаем матрицы моделей для случайного сигнала, если исходные данные будут случайными, то вот эти все величины, количество информации, хи-квадрат, они будут стремиться к нулю, потому что как раз хи-квадрат - это критерий зависимости принадлежности объекта к классу от того, что у него есть определённые признаки. То же самое количество информации, как я вам объяснил. И рой коэффициент возврата инвестиций, тоже то же самое по смыслу. Я вам это всё показал. Получается, что если мы возьмём сигнал, состоящий из шума и полезного сигнала, то шум приведёт к нулевому вкладу при увеличении объёма выборки. И полезный сигнал мы вытащим и будем использовать для решения задач. А в факторном анализе параметрический метод говорится так, что факторы должны быть независимы, ортонормированы должны быть. И факторов должно быть не больше семи, а у нас уже здесь вот их 11. И абсолютно, исходные данные должны быть абсолютно точными, потому что даже небольшое их изменение влияет существенно на результаты решения задач. Ну если это так, ребята, то тогда факторный анализ вообще неприменим для вас. Но вы его будете применять, потому что вам скажут: "Применить факторный анализ". Вот все применяли, и вы примените. Но это будет некорректно. Значит, есть такой выдающийся математик российский, современный. Я именно без кавычек это говорю. Орлов Александр Иванович. По-моему, я вам про него говорил. Вот, он наиболее цитируемый современный математик России, ребята. Доктор экономических наук по специальности 082013, доктор технических наук, кандидат физмат наук, профессор Орлов. Ну есть работа эконометрика, основы эконометрики. Я вам очень советую посмотреть эту работу, хотя вы не экономисты, и вроде бы как вам причём здесь эконометрика? При том, что он описывает, какие условия корректного применения методов параметрической статистики. Просто перечисляет. И выясняется, я прочитал эту работу. У него прекрасный язык, кстати, он пишет научные работы так, что они читаются как художественное произведение, легко, увлекательно, с какими-то примерами интересными там, ну, чувство юмора там так сдержано, но, может быть, и можно заметить. У меня бывает прямо ярко выражено. Вот, в работах, я иногда не стесняюсь даже. Вот. Так вот, по этому поводу. Так вот, у него, я когда прочитал эту его работу, у меня возникло устойчивое такое впечатление, что вообще невозможно корректно применить эти методы. Что это сложно, понимаете? Что люди, которые этим занимаются применением реальным, они даже не представляют себе, за что они вообще взялись. Они не доказывают гипотезу о выполнении нормального распределения, применяют параметрические методы, не проверяют, насколько связаны факторы, то есть не кластерный анализ факторов не проводят. Применяют методы, где написано, что они не должны зависеть друг от друга, а они зависят друг от друга. Они это не проверяют. И вот так вот подряд, понимаете? То есть не одна ошибка, а у них десятки ошибок методических. Одна на другую наслаивается, и диссертация защищается. То есть, ну вообще кошмар какой-то, понимаете? То есть я вам прямо серьёзно это говорю, что корректно это применить очень сложно. И поэтому я вот эти методы и разрабатывал и говорю о том, что они применимы вот в этих случаях. Это проверено, это доказано, что они применимы.

10. Завершение

Вот. Следующее, что мы можем рассмотреть. Значит, вопрос возникает такой: среди этих значений факторов, которые мы вот здесь вот у нас есть эти значения факторов, все ли имеют одинаковый вес и ценность для решения тех задач, которые мы решали в системе? И мы сейчас возвращаемся к вопросу, в статье этот вопрос тоже рассматривается, степень детерминированности будущих состояний и ценность факторов для решения задач. И мы смотрим, что оказывается, у нас факторы не все имеют одинаковую ценность. Вот сейчас мы берём модель, заходим в F3. Ну здесь, к сожалению, не так уж это заметно. Вот. Ну, скажу так: 50% наиболее значимых значений факторов обеспечивает 61% суммарной значимости, а 40% наиболее значимых факторов обеспечивает 50%. Ну здесь вот у нас больше различия. 50% обеспечивает 70%. Опять же мы видим по кривую. Вот. То есть она вот так вот идёт выгнута, потом близко к параллельной. То есть сначала расположены наиболее ценные значения факторов, потом менее ценные. И то же самое касается степени детерминированности будущих состояний объекта моделирования. Всё это есть в графической и в табличной форме. Степень детерминированности. Насколько какие будущие состояния, насколько жёстко они обусловлены действующими факторами. Есть некоторые состояния, которые прямо жёстко обусловлены, однозначно. А есть, которые слабо обусловлены.

Вот. Ну здесь вот, допустим, 50% наиболее значимых, наиболее жёстко обусловленных классов обеспечивает около 80% суммарной детерминированности, обусловленности. Ну то есть тоже есть классы, которые более жёстко обусловлены, и которые менее жёстко обусловлены. Те, которые менее жёстко обусловлены, они достигаются с меньшей гарантией, чем жёстко обусловленные. То есть жёстко обусловленное - это что? Фактор действует, состояние получено. Слабо обусловлено - факт, много факторов нужно, чтобы действовали, тогда состояние будет получено. Если не очень, не все факторы соответствуют этому, то тогда, возможно, и не будет оно получено это состояние. То есть мы знаем это всё количественно, видим и в табличных формах, и графических. Смотрим теперь на интегральный критерий другой - семантический резонанс знаний. Он очень напоминает первый интегральный критерий - сумма знаний, но только здесь мы видим, что вектора стандартизированы. Вектор классов стандартизирован. То есть из каждой значения координаты вычтено среднее по вектору от координат, делится на среднеквадратичное отклонение. И умножаются такие вот стандартизированные уже вектора друг на друга. Выносятся общие сомножители за сумму. Вот. И мы видим, что это тоже скалярное произведение, только оно скалярное произведение отнормированных уже, стандартизированных векторов. Можно по-другому их стандартизировать. Можно взять из значения координаты, вычесть минимальное и разделить его на максимальное минус минимальное. Тоже будет нормировка от нуля до единицы. Вот. Но это такая более традиционная, а то и меньше известна в линейной интерполяции и нормирование к этому диапазону 0-1. Это более известная нормировка, поэтому я её использую. Ну они вообще мало отличаются по смыслу. Если мы посмотрим на эту формулу, то мы можем узнать, что это коэффициент корреляции Пирсона. То есть Пирсон, он посчитал, что два массива сходны или отличаются друг от друга на эту вот величину, которую можно таким образом посчитать. И это, ребята, является обоснованным, хотя он об этом не писал, но это является межвекторным расстоянием, по сути дела, где только вектора стандартизированы, которые сравниваются. То есть это скалярное произведение двух стандартизированных векторов. Сейчас я вам сошлюсь на книжку, где об этом говорится. И эту книжку вы легко можете найти в интернете и посмотреть. Вот там написано.

Вот. Теперь, ребята, мы так более-менее подробно рассмотрели пример, который я хотел вам показать, о том, который показывает, как решается задача прогнозирования, принятия решений и исследования моделируемой предметной области на задаче, которая имеет отношение к вашей специальности. Ну не совсем это точно то же самое, но что-то такое близкое, то, что вы понимаете, то, что вы проходили, изучали. И лучше вам понятно это. Вот. Значит, теперь что я могу ещё вам предложить? На следующем занятии мы можем исследовать графические изображения. То есть можем исследовать спектральный анализ провести изображений листьев, например, поражённых различными вредителями или разная степень повреждения какой-то болезнью. Можем рассмотреть, как их отличить друг от друга. То есть как можно создать измерительный инструмент, позволяющий измерить степень поражения листа каким-то вредителем или растения. Если есть интерес к этому. Вот. Вообще, у нас так получается, что вроде бы по учебному пособию, если смотреть, то мы прошли пятую лекцию, изуч прошли. Вот. У нас их там, по-моему, шесть. Следующая лекция - это она посвящена, если пособие взять, посмотреть, она посвящена перспективам применения этих технологий в области садоводства. Значит, я бы очень вам рекомендовал, ребята, прямо вот, не знаю, настоятельно. Значит, возьмите какие-то свои данные, подготовьте по образцу, который я вам показал. Ну если у вас будут какие-то ошибки, мы их прямо на лету исправим быстро эти ошибки и сможем, значит, показать этот образец, где он? Этот образец. Значит, давайте, знаете, что сделаете? Вот староста, пришлите мне письмо, я вам его пошлю этот образец сейчас. И вы наподобие можете сделать свои образцы, то есть какие-то свои данные. Строчки - это наблюдения, колоночки - это результаты в стоимостном выражении и в натуральном выражении. В стоимостном - прибыль и рентабельность, в натуральном - количество и качество продукции. И какие-то факторы, которые влияют на получение этих результатов: природно-климатические факторы и факторы агротехнологии, а также факторы свойства самого сорта, которые тоже очень важными являются. То есть здесь это представлено в виде морфологических свойств сорта, гибрида. Понятно, да, ребята? И мы всё это обработаем. То есть факторы делятся на природно-климатические, факторы управления, агротехнологии и факторы, которые связаны со свойствами самого сорта. Я это показывал. Это свойства самого объекта управления, текущие и предыдущие, предшествующие. Ну, допустим, предшественники какие были там.

Сейчас, конец занятия. Всего самого хорошего. До свидания. На следующем занятии решаем ваши задачи.

До свидания.

До свидания. Спасибо. До свидания. Спасибо. До свидания.

До свидания. На ваших данных. Значит, где адрес? Ага, вот этот адрес. Это я адрес я, конечно, писал уже. Сейчас ещё раз напишу. Вот. Не отходя от кассы.

До свидания. Прекращается запись. Я пошлю сейчас, Алин.