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

***Российская Федерация***

**199 Работа с Эйдос 2020-11-11**

**Заголовок:**
Применение системы Эйдос для анализа данных: спектральный и статистический анализ

**Резюме текста (1 страница):**

**1. Введение и организационные моменты**

* Лекция начинается с приветствия и обсуждения технических проблем со звуком у лектора.
* Отмечается малое количество присутствующих студентов, предполагается связь с проходившим ранее тестированием.

**2. Обзор предыдущих тем и знакомство со специализациями студентов**

* Лектор просит напомнить тему прошлой лекции (АСК-анализ в системе Эйдос) и предлагает продолжить.
* Студенты представляются, называя свои специализации:
	+ Общее земледелие и растениеводство (сельское хозяйство).
	+ Плодоводство и виноградарство (сельское хозяйство).
	+ Биологические науки (биотехнологии, включая бионанотехнологии).
	+ Промышленная экология и биотехнологии (разработка консервных изделий функционального назначения).
	+ Экология и природопользование (экология насекомых).
	+ Промышленная экология и биотехнологии (комплексная переработка виноградных выжимок, функциональные продукты питания).
* Лектор предлагает адаптировать лекцию под интересы студентов или дать задание на разработку собственного приложения в Эйдос для обработки их данных.

**3. Демонстрация 1: Спектральный анализ изображений (повреждение листьев)**

* Лектор решает показать пример разработки собственного приложения в Эйдос на основе данных из интернета.
* **Задача:** Оценка степени повреждения (пятнистости) листьев ячменя.
* **Метод:** Спектральный анализ изображений листьев с разной степенью повреждения (от 0% до 80%).
	+ Загрузка изображений в систему Эйдос.
	+ Генерация цветовых спектров для каждого изображения (представление доли каждого цвета/оттенка). Система делит спектр на заданное число (здесь 35) частей.
	+ Создание обобщенных спектральных образов (моделей) для каждого класса повреждений. Сравнение спектра конкретного листа со средним спектром по всей выборке (пунктир) для выявления характерных и нехарактерных цветов для данного класса.
	+ Интерпретация: Для здоровых листьев характерны зеленые тона, для сильно поврежденных – коричневые.
	+ Анализ информативности цветов: Выясняется, какие цвета (спектральные диапазоны) несут больше всего информации для различения степени повреждения.
	+ Проверка достоверности моделей и визуализация сходства/различия классов с помощью дендрограммы и 2D-диаграммы. Результаты показывают, что система может группировать листья по степени повреждения на основе их цветовых спектров, но есть сомнения в корректности экспертной оценки некоторых примеров (например, 30% повреждения группируются со слабо поврежденными).

**4. Демонстрация 2: Статистический анализ табличных данных (Гумус и урожайность)**

* **Задача:** Выяснить, как содержание гумуса в почве влияет на урожайность различных сельскохозяйственных культур.
* **Данные:** Табличные данные (предположительно, из статистики или научной статьи) с содержанием гумуса и урожайностью (люпина, картофеля, овса, ржи, ячменя).
* **Метод:** АСК-анализ в системе Эйдос.
	+ Ввод данных: Таблица с колонками "Гумус" (фактор) и урожайностью культур (результаты).
	+ Кодирование данных: Непрерывные значения (содержание гумуса, урожайность) разбиваются на диапазоны (например, низкое, среднее, высокое содержание гумуса/урожайность).
	+ Построение моделей: Система строит модели, показывающие связь между уровнем гумуса и уровнем урожайности каждой культуры.
	+ Интерпретация моделей (когнитивные функции): Визуализация связей показывает, что низкое содержание гумуса жестко обуславливает низкую урожайность люпина. Высокое содержание гумуса сильно влияет на высокую урожайность картофеля и ржи. Связи для других культур менее выражены или требуют дальнейшего анализа.

**5. Прочие примеры и заключение**

* Лектор упоминает возможность анализа связи экологических показателей с развитием промышленности на основе данных Росстата.
* Также упоминается анализ контуров листьев винограда для идентификации сортов.
* Подчеркивается возможность применения системы Эйдос для решения разнообразных задач анализа данных в областях, релевантных для студентов.
* Лекция завершается предложением проанализировать данные самих студентов в будущем.

**Детальная расшифровка текста:**

**Раздел 1: Введение и организационные моменты**

**Подраздел 1.1: Приветствие и технические проблемы**

Здравствуйте.
Здравствуйте.
Что-то вас мало.
У вас плохо слышно.
Говорю, что вас мало, говорю.
Сейчас слышно, нет?
Слышно, но какое-то шипение, чириканье.
Слышимость.
М-м.

**Подраздел 1.2: Обсуждение количества присутствующих**

Ладно.
Мам, почему так мало вас, не знаете?
Может быть, сейчас ещё подойдут, ребята, ну, точнее, зайдут.
Угу. Вы были на тестировании сегодня, нет?
Да, ходили.
Угу. Может, с этим как-то связано?
Ну так тестирование у нас было с 10:00 до 11:00.
М-м.

**(Пауза, лектор работает с программой)**

М-м.
Понятно.

**(Длительная пауза, лектор работает с программой, звуки кликов мыши)**

**Раздел 2: Обзор предыдущих тем и знакомство со специализациями студентов**

**Подраздел 2.1: Напоминание о прошлой теме**

Ладно, ребята.
Давайте тогда, напоминайте, на чём мы остановились, и продолжим.

**(Пауза)**

Вы можете вспомнить, на чём мы остановились, нет?

**(Пауза)**

Ну точно вспомнить не сможем, но помним, что вы Эйдос показывали.
Да.

**(Пауза, звук подключения нового участника)**

Ну ладно. Я просто смотрю, я уже три лекции подряд одно и то же записываюсь.
Травка об АСК-анализе в системе Эйдос у меня записано.

**(Пауза, лектор работает с программой)**

**Подраздел 2.2: Предложение адаптировать лекцию и знакомство со студентами**

Так, ладно.
Может быть, у вас какие-то вопросы есть? Или у вас какие-то есть пожелания? Чтобы я вам рассказал о том, что вас интересует больше?

**(Пауза, шум на фоне)**

Всё себе представляете, да?

**(Пауза, лектор работает с программой)**

Ладно, ребят. Значит тогда продолжим, наверное.
Я вам какое-то задание давал, нет? Говорил, что имеет смысл сделать своё приложение в системе? Ваши данные какие-то обработать, которые у вас есть?
Такие делал предложения вам?

**(Пауза)**

Ладно, ребята, тогда я сейчас вам буду показывать, как можно разработать собственное приложение на основе данных, которые можно найти в интернете.

**(Пауза)**

Кстати, что интересно, вот у меня аспирантских занятий вообще здесь нету, видите?
У меня вот вашей вашей пары нету.

**(Пауза, лектор работает с программой)**

**(Пауза)**

Ребят, сейчас, пожалуйста, скажите, в какой области у вас работа? Какие у вас специальности?
Вот Александра, Амина, Анастасия, Мира.
И, наверное, Астамур и Альбина.
Вот у нас такая компания.
Евгений Вениаминович, у меня общее земледелие, растениеводство, специальность, направление сельское хозяйство.
Угу.
У Жакировой сельское хозяйство, плодоводство, виноградарство.
Угу. Понятно.
У Забашты биологические науки, биотехнология, в том числе бионанотехнология.
Круто. А бывает ещё? Ну ладно. Ну, а дальше? Амина?
Акоба, а промышленная экология и биотехнология. Тема связана с разработкой консервных изделий функционального назначения.
Ничего себе. А что такое функциональное назначение?

**(Пауза)**

Консервные изделия функционального назначения.
Абдрахманова, у меня экология природопользования, а тема связана с экологией насекомых.
Угу. Прикольно.
Малеева Альбина. Промышленная экология и биотехнология. Тема - разработка технологии комплексной переработки виноградных выжимок с получением функциональных продуктов питания.

**(Пауза)**

Интересно.
Спасибо.
Вообще звучит так серьёзно всё у вас.
Вот. Ну что?
У вас есть какие-нибудь данные собственные, чтобы вот на ваших данных попробовать какой-то провести исследование небольшое?
Я могу вам привести пример, э-э, просто вот из своих работ. Допустим, показать вам работу, связанную с экологией, с исследованием экологии. Или могу показать работу, связанную с исследованием, э, допустим, м-м, в области ампелографии, например. Что вам интереснее?
Ну давай про ампелографию расскажу вам. То, что мы делали с Леонидом Петровичем.

**(Длительная пауза)**

Мне бы хотелось как-то вам показать, что можно было бы использовать эти вот технологии, которые я разработал в вашей области.

**(Пауза, лектор печатает)**

Вам показывал работу по спектральному анализу, нет, ребят? Изображений?
Нет, не показывали.
Не показывал, да? Давайте сейчас покажу. Это будет интересно.

**Раздел 3: Демонстрация 1: Спектральный анализ изображений (повреждение листьев)**

**Подраздел 3.1: Постановка задачи и ввод данных**

Ну с чего можно начать? С того, что нужно какие-то изображения взять, которые мы будем анализировать.

**(Пауза, лектор работает с программой, печатает)**

Вот. Ну, допустим...

**(Пауза, лектор работает с программой)**

Вот смотрите, ребята. Значит, мы можем взять вот эту инструкцию. И по этой инструкции попробовать поработать. Послал в чат вам ссылочку на эту инструкцию.
И вот, к примеру...

**(Пауза, лектор работает с программой)**

Так, у нас...

**(Пауза, лектор работает с программой)**

Вот, ребят, вот смотрите, вот, значит, есть листики с разной степенью повреждения, видите, да? Пятнистостью. Вам это должно быть понятно всё.
И вот возникает вопрос, как измерить степень повреждения? Обычно это делается с помощью каких-то таблиц, атласов, сравнением, значит, на взгляд это, так сказать, как это выглядит. Ну, вручную, грубо говоря, это делается, просто смотрят на них и всё.
А вот, значит, как это можно сделать в системе Эйдос. Значит, тут мы сейчас попробуем это сделать.
Значит, в системе Эйдос есть много интерфейсов программных, ну это такие средства, режимы, позволяющие вводить изображения, э-э, тексты и табличные данные. Причём табличные данные могут быть и числовые, и текстовые.

**(Пауза, лектор работает с программой)**

Сейчас мы вводим эти изображения в систему Эйдос.

**Подраздел 3.2: Анализ спектров и создание моделей**

И потом начнём их анализировать по их спектрам. То есть система сейчас измерит спектры этих изображений. Но самое главное, что не просто спектры измерит, а ещё создаст обобщённые спектры категорий различных.

**(Пауза, лектор работает с программой)**

То есть мы должны сейчас всё это видеть, да, ребят? То есть видите изображение?
Да, всё видно.
Угу.

**(Пауза, лектор работает с программой)**

Это реальная задача.

**(Пауза, лектор работает с программой)**

То есть я вам просто хочу вам показать, как можно решать задачи различные.

**(Пауза, лектор работает с программой)**

По экологии, ребята, у меня интересные есть работы. У меня была мысль такая: как связана экологические показатели с развитием промышленности? Я взял... пока там оно щёлкает, я вам покажу.
Я взял со стата статистики федеральной по регионам России состояние экологии, экологические показатели. И тоже с этого же сайта развитие промышленности.

**(Пауза, лектор работает с программой)**

У меня получились интересные вещи. Оказалось, что есть прямая связь. Ну, оно как бы и так интуитивно понятно. Вот сейчас я в чат направлю ссылочки на некоторые свои работы по экологии. Их довольно много.

**(Пауза, лектор работает с программой, печатает)**

Вот здесь вот нашлось 10.

**(Пауза, лектор работает с программой, печатает)**

Вот это по экологии, некоторые статьи по экологии.

**Подраздел 3.3: Интерпретация результатов спектрального анализа**

Теперь смотрите, ребята, что написала система Эйдос? Что теперь нужно зайти в режим 2 3 2 3 для ввода, завершения ввода этих изображений. 2 3 2 3.
Всё, изображения введены.
Мы видим у нас, ребята, смотрите, у нас 11 есть э-э категорий изображений. Это ячмень э-э с различными сте- степенью поражениями э-э, с различной степенью поражения площади листа пятнистостью. От нуля до 80%. Это по экспертным оценкам, то есть это так, как считают специалисты.
И у нас э-э есть э-э спектр, который разделён на 35 частей. Количество частей, на которые мы делим спектр, задаётся нами. Ну, там вот э-э, когда я эту задачу задавал, запускал режим, там было задано 35. Ну, из моего опыта, я могу сказать так, что больше будет сливаться эти цвета, не будет их, ну, на глаз их будет сложно отличить. Если взять их меньше, то будет грубовато. Если больше, то особого смысла нет, грубо говоря.
И есть обучающая выборка. Значит, здесь вот у нас эти листики, пока для каждого класса один пример только. А здесь у нас коды спектральных диапазонов, которые встречаются в этом листике. То есть вот эти вот здесь вот коды этих спектральных диапазонов.
Вот. Дальше мы должны создать модель и проверить их на достоверность эти модели.

**(Пауза, лектор работает с программой)**

Это всё описано в статьях, ребята. Сейчас я вам найду статьи э-э по спектральному анализу изображений. 2017 год.
Модели созданы. Мы можем посмотреть, какая у них достоверность попытаться.

**(Пауза, лектор работает с программой)**

Вот. Теперь можем посмотреть, как решаются различные задачи. Выберем наиболее достоверную модель. Ну, здесь модели получились не очень высокой достоверности. Ну сейчас попробуем сделать получше. 3 5

**(Пауза, лектор работает с программой, печатает)**

Сейчас модели создаются и проверяются на достоверность при разных параметрах.

**(Пауза, лектор работает с программой)**

Значит, что здесь такое? Значит, мы здесь видим, ребят, здесь вот у нас уровень сходства конкретного листика с обобщённым образом. Он может быть положительный и отрицательный. Значит, положительный, когда листик похож на обобщённый образ, а отрицательный, когда не похож. И, значит, здесь уровень сходства. Красное - это количество таких э-э случаев, когда, ну, в общем, называется это истинно положительное решение, когда система правильно отнесла листик к той категории, к которой он относится. А категория по степени поражения. А синяя - когда неправильно. То есть есть и ошибочные решения, но при очень низких уровнях сходства, когда до 15%. Выше там э-э всегда, как правило, истинных решений больше.
Теперь посмотрим, как она создала сами эти спектральные образы. Сначала смотрим на... Здесь она пишет, что будет создана папочка, вернее, уже создана, куда будет всё это записываться.
Значит, вот мы видим э-э какие у нас э-э изображения. Это по несколь- по несколько листиков у нас есть, по несколько примеров по разным степеням повреждения.

**(Звонок телефона, лектор отвечает)**
Да? Здравствуйте, Таня. А, Жень, здравствуйте, Таня, не отвлекаю? Отвлекаешь. Да, ну на секундочку. Я хотела спросить, как там Глеб... У меня занятие сейчас. А опять? На удалёнке? Ну да, да. Да, я тебе пере- перезвоню. Ага, ну-ну. А Глеб нормально? Да, нормально. Ну слава Богу, переживаю. Слава Богу. Потом перезвоню. Угу. Давай. Люсе скажу, что ты звонила.

**(Лектор продолжает лекцию)**

Вот. Значит, что мы здесь видим? Значит, мы видим э-э изображения листика. На нём есть пятнышки коричневые. А здесь спектр, какой э-э цвет э-э пикселей насколько представлен в этом изображении. И в процентном отношении. Вот здесь вот 100% этих пикселей берётся и э-э из них там 5% зелёных, там 4% коричневых. Короче говоря, мы видим спектр каждого листика.
Вот. Точно так же мы можем измерять э-э, определять спектр, ну, цвет фактически, очень точно определять цвет различных объектов, которые сфотографированы и фотографии их изображения помещены в папочку для анализа системы.
Ну, э-э, каждый э-э конкретный объект э-э проанализирован с точки зрения того, какие цвета, в какой степени там представлены. Объект выборки обучающий.
А теперь давайте посмотрим на сравнение обобщённые образы классов, когда конкретное изображение сравнивается с обобщёнными. И здесь мы видим вот сами эти обобщённые образы. Сейчас я вам объясню, что здесь такое мы видим.
Значит, смотрите, значит, вы, наверное, замеча- заметили, что здесь есть пунктир. Ну, правда, возможно, он вам виден не очень хорошо, но мне хорошо видно. Пунктир вот здесь. Видите, да? Ребят? Пунктир. То есть есть спектр, и есть пунктир.
Ребят, вы меня слышите вообще, нет? Ребята? Вы там?
Да-да, слышим.
Что-то у меня возникло некоторое сомнение смутное, что вы там.
Значит, смотрите, значит, если мы возьмём все листики, вот которые у нас там были в качестве примеров приведены, и будем их рассматривать как одно изображение, вот все вместе они. То тогда там э-э каждый цвет представлен в какой-то степени. И вот этот пунктир, он отражает э-э степень э-э насколько этот цвет представлен во всей совокупности вот этих вот изображений.
Вот мы видим, допустим, что зелёного там побольше, чем здесь вот, вот видите, пунктир? А коричневого там такого зелёно- зелёно-желтоватого, коричневого поменьше, чем в этом конкретном изображении.
И вот э-э мы знаем, как представлен цвет, каждый каждый цвет в среднем по всей выборке изображений. А в данном конкретном изображении повреждение на 80%. Вот здесь вот у нас вот эта коричневая гамма, коричнево-зелёная гамма представлена сильнее, чем в среднем по всей выборке. А вот эти вот зелёные цвета тоже там есть, но они представлены меньше, чем по всей выборке, относительно меньше.
Ну это примерно, я вот когда студентам рассказываю про эти задачи, то я примерно э-э так рассказываю. Вот представьте себе, что нам нужно определить, кто-то за дверью стоит, студент или студентка. Нам нужно определить, кто там стоит. И мы можем задавать вопросы, и нам могут отвечать на них да или нет. Вот я спрашиваю: у вас волосы длинные? Мне говорят: да. Значит, я думаю, наверное, студентка. Почему я так думаю? Вот почему у меня такая мысль появляется? Потому что длинные волосы в группе студентки или аспирантки встречаются значительно чаще, чем в среднем. Почему чаще, чем в среднем встречается? Потому что в среднем ещё кроме студенток, ещё есть студенты. Кроме аспиранток, ещё есть аспиранты. А у аспирантов обычно волосы короткие.
Вот. Поэтому это этот признак, длинные волосы, он характерен для группы студентки, девушки, и не характерен для группы ребята. Понятно, да? То есть у ребят он встречается реже, чем в среднем. Встречаются ли у ребят длинные волосы? Встречаются, бывают, но реже, чем в среднем, причём намного реже, чем в среднем. А у девушек они встречаются намного чаще, чем в среднем.
Вот. То есть мы получаем э-э такую информацию фактически. Если у нас какой-то признак встретился, то этот признак несёт информацию о том, что объект относится к определённому классу, определённой категории, или или о том, что он не относится к определённому классу. Ну, скажем, если встретился признак длинные волосы, то это большой объём информации о том, что это студентка.
Вот представьте себе, что этот класс обобщённый, образ этого спектральный, этого обобщённого класса повреждения 80%. Это вот эта гамма коричнево-зеленоватая такая. А ярко-зелёная отсутствует там, или не отсутствует, но встречается гораздо реже, чем в среднем. Вот если такой цвет встречается, значит, это, скорее всего, э-э повреждённый листик. То есть это несёт информацию о том, что лист повреждён.
И вот мы можем посмотреть эти все изображения записаны в папочке соответствующей. Вот мы видим э-э для низкой степени повреждения, видите? Характерно зелёные тона, а коричневая гамма встречается, но но гораздо реже, чем в среднем. И вот так вот мы можем смотреть. Так, это у нас что-то она как-то прыгает нелогично. Сейчас попробуем сделать, чтобы логично было. Вот. А, ну да, здесь нумерация такая. Ну значит, так и будем показывать.
Вот, значит, нулевое повреждение, зелёненький, короче он говоря, видите? А это вот э-э большое повреждение, но не самое большое, но большое. 70%. Это самое большое повреждение. А это почти не повреждённый. Видите, всего 5% повреждений. Это чуть больше. Это ещё больше. Ещё больше повреждений. Ещё больше, ещё больше.
Ну я могу сказать, что когда вот мы эти изображения анализируем, то у меня возникает вопрос о том, насколько корректно отнесены экспертами, э-э эксперты, насколько корректно эксперты оценили степень повреждения. Потому что по системе, если смотреть, то немножко не сходится.
Ну теперь давайте посмотрим насчёт того, насколько эти листики, то есть э-э обобщённые образы листи- листьев с разной степенью повреждения сходны друг с другом. Вот сейчас мы посчитали матрицу сходства этих листьев. И сейчас получим диаграмму, которая отражает степень их сходства, различия.
Вот мы видим, что э-э очень низкое повреждение 1 11, сходно с 2 11. Вот, э-э, а 2 11 сходно 6 11. Это сходно 3 11. Вот группа. Вот мы можем сказать так, а потом ещё объединяются 4 и 11 1 11. То есть что это такое? Это вот красное - это сильно повреждённые листики. А вот эти вот синенькие - это слабо, вообще не повреждённые 00. И повреждённые на 5%, на 10%, вот они сходны друг с другом. И сюда же к ним примыкает на 30%. То есть я вот э-э не специалист по фитопатологии, но если вот посмотреть на эту дендрограмму, то у меня возникает такое подозрение, что вот эти листики, которые отнесены к тридцатипроцентным повреждениям, они туда отнесены ошибочно. То есть, скорее всего, у них м-м ну они должны быть отнесены к другой категории по степени повреждения, к более низкому повреждению. Потому что они оказались похожими вот на те, кто на 5 и на 10% повреждён. А на 15 оказались в другой группе вообще, видите?
Вот. То есть 15 они более повреждены, чем 15, а 30 повреждены менее, чем 30. Вот у меня такое подозрение. В общем, здесь не совсем правильно э-э экспертами было оценена принадлежность этих листьев к этим классам по степени повреждения.
Теперь смотрим, что характерно у нас для тех или иных образов. Вот берём, допустим, листья, которые у нас зелёненькие. Это нулевое повреждение. И смотрим, что для них характерно? Для них характерно вот эти цвета, которые слева. Зелёные и чуть-чуть желтоватые, коричневатые. А вот эти, которые справа, коричневые, э-э, и всякие красные, пурпурные, они вообще не характерны для них. Вот смотрим это в графической форме. Значит, класс ячмень 0% повреждения. Слева то, что характерно, а вот эта линия, толщина линии означает степень характерности. Больше всего характерно вот такой спектральный диапазон цветовой. 10 35 из этого спектра всего. То есть это зелёненькая, зелёный цвет вот такой.
А вот такой коричневый и эти, этот зелёный тоже почему-то не характерен, тринадцатый. Вот. Ну, так вот оказалось это.
Вот. А если мы возьмём наиболее сильно повреждённые, то там картина такая, что здесь вот мы видим, какие цвета наиболее характерны для этих наиболее сильно повреждённых листьев - коричневые, желтовато-коричневые, даже вот такие синие есть. А то, что не характерно для них - это зелёные цвета не характерны.
Вот. И мы можем узнать, какой цвет, какое количество информации несёт о принадлежности листика с этим цветом к тому или иному классу. Вот, допустим, берём вот такой цвет красный. И как возникает вопрос, это о чём говорит? О том, что он повреждённый или нет? Как вы считаете? Ну это интуитивно понятно, что это повреждённый лист. Вот его цвет красный, такой алый. Повреждение больше всего похоже, что 40%, 30%, 20%, 50%, 70 и 60. Они похожи, чтобы 0% было повреждения, 5%, 10%. И 80 тоже не похоже. 15 и 10. Вот явно здесь видно, что некоторые листики там оценены были, видимо, не совсем правильно. Этот цвет вообще очень редко встречается, он вообще не встречается практически. Вот.

**(Пауза, лектор работает с программой)**

Вот такой спектральный диапазон оранжевый характерен для средней степени повреждения листьев. То есть мы можем сразу понять, какой смысл у того или иного цвета.

**(Пауза, лектор работает с программой)**

**Раздел 4: Демонстрация 2: Статистический анализ табличных данных (Гумус и урожайность)**

Вот. И получается некий массив, который можно дальше исследовать. Ну а с контурами мы не будем сейчас смотреть.

**(Пауза, лектор работает с программой, печатает)**

Так, ну мысли возникли, ребят? По поводу того, можно ли это как-то использовать?
И э-э давайте сейчас попробуем найти какую-нибудь задачу, которая ближе к вашим темам.

**(Пауза)**

У кого в теме есть слово "влияние"? Влияние там чего-то на что-то?

**(Пауза)**

Или смысл такой, что влияние чего-то на что-то?

**(Пауза, лектор работает с программой, печатает)**

Ребят, вот смотрите, видите? Значит, я попробовал найти таблицу какую-нибудь, которая бы содержала информацию о том, как связано содержание гумуса, например, в почве с урожайностью разных культур. И, значит, э-э разным видам культур. Вот. Сейчас мы эту таблицу перетащим в Excel и попробуем эту задачу решить в системе.
Значит, в системе э-э, значит, э-э нужно знать, что наименование э-э колонок должны быть в одной строке. Поэтому я разбил таблицу сразу же на части и сделал одну строку заголовков. Второе, что сделал, сейчас я сделаю это. Я возьму эту таблицу, скопирую в другой лист, вставлю, а тот лист удалю. Для чего? Чтобы у меня область данных была точно соответствовала фактическому их расположению. Потом добавлю колоночку и напишу здесь вот номер. Здесь вот сделаю номер наблюдения. Ну, на самом деле, здесь среднее наблюдение, но будем считать, что это у нас наблюдение. Ну, в принципе, это могло бы быть поле, например.
Вот. Теперь смотрите, ребята, значит, вы когда будете делать таблицы у себя там в работах, вы имеете в виду, что надо, чтобы в одной в колоночке число знаков после запятой э-э было одно и то же. Вот, допустим, у вас здесь два знака после запятой, а здесь вот один знак после запятой. Видите, да, ребята? Что это значит? Это значит, что у вас э-э один знак после запятой правильный, а остальное вы не знаете. А если вы их не знаете, зачем тогда писать вот эти вот знаки, два знака после запятой? То есть отсюда вывод такой, что либо надо здесь вот написать 1 7 0, либо вот здесь везде писать, округлять 1 2 1 4 1 6 там и так далее.
Вот. Ну я думаю, что лучше написать два знака после запятой. Это к чему я говорю? К тому, что элементарная такая грамотность должна у вас быть. Вот, число, два знака после запятой. Потом берём формат и вот так вот раз. И вот здесь тоже. А здесь а здесь и не надо было. Здесь и не надо. А, нет, здесь надо, потому что здесь указано до одного знака после запятой, а не два.
Вот. Ну сейчас табличка корректная. Такая табличка ни у кого не вызовет проблем.
Теперь смотрите, ребята, у нас содержание гумуса является фактором, который влияет на урожайность разных культур. То есть урожайность у нас - это результат влияния фактора. То есть результаты влияния мы изображаем жёлтым фоном и считаем колоночки здесь. Значит, у нас Shift держим, в клеточке А1 я держу, э-э, ставлю курсор, ставлю, держу Shift и двигаю вправо. Вторая колоночка - это у нас описательная шкала. А с третьей по восьмую - классификационные шкалы. Вот, и мы это дело запишем.

**(Пауза, лектор работает с программой)**

И мы посмотрим, как влияет содержание гумуса в почве на урожайность различных культур. Можно было бы ещё добавить сюда вот этих данных из других таблиц. Ну пока что так вот будем обойдёмся таким простым примером.
И попробуем этот пример обработать, что у нас получится. Сначала введём данные в систему. Чтобы это сделать, для этого мы запускаем режим 2 3 2 2. И указываем тип файла экселевского. Потом указываем, что у нас э-э классификационные шкалы с результатами, с классификацией результатов с третьей по восьмую. А описательная шкала у нас вторая, то есть это фактор, только вторая колоночка. И указываем, что у нас э-э диапазоны числовые в этих шкалах брать разного размера, но с одинаковым числом наблюдений. Для чего это нужно? Это связано с тем, что у нас э-э диапазон изменения величины в колоночке э-э какой-то существует в каждой колоночке. И данные распределены в этом диапазоне неравномерно. Ну, допустим, все тяготеют к минимальным значениям или к максимальным значениям.
Вот. Ну сейчас мы посмотрим, как у нас они распределены эти данные.

**(Пауза, лектор работает с программой)**

Ну сейчас вроде бы получилось. Вот смотрите. Значит, классификационная шкала у нас три наблюдения на числовой диапазон. Описательная шкала - в одном диапазоне даже два наблюдения только, потом три. Ну, табличка маленькая, данных мало, поэтому это естественно. Созданы справочники классификационные, описательные шкалы градаций. То есть это справочники, с помощью которых кодируются результаты и кодируются э-э факторы. Вот, зелёная масса люпина, картофель. Как они получены? Найдено минимальное значение и максимальное в колоночке, разделено на три числовых диапазона разного размера, сами диапазоны, но такие, что у них наблюдается одно и то же число наблюдений. И фактор. Фактор у нас один - гумус. Ну, могло бы быть не только гумус, могло бы быть ещё там удобрения, средства защиты, полив, вспашка. Это просто у нас в этом примере только гумус. И обучающая выборка. Что мы здесь видим? У нас действовал первое э-э значение, код код один из шкалы факторов, и получилось у нас вот такие результаты. Это коды результатов. И всего у нас восемь примеров. Создаём модели. Проверяем их на достоверность. Достоверность получилась очень высокая у всех моделей. 0 0,947 при максимуме единица. Вот, это выше там некуда уже.
Вот. И мы видим, что, значит, ложные истинные решения, истинных решений больше, начиная с 60%, ложных решений не наблюдается. Ну, начиная вот, да, вот с этого уровня, только истинные решения. Можно здесь написать название исследования. АСК-анализ влияния содержания гумуса в почве на урожайность различных сельхозкультур.

**(Пауза, лектор работает с программой, печатает)**

Вот. И посмотрим, что у нас получилось. 4 4. Да, во-первых, какие у нас модели получились? Давай посмотрим. Вот у нас э-э результаты э-э по содержанию гумуса - низкое, среднее, высокое. И результаты э-э зелёная масса люпина, зелёная, значит, низкая, средняя, высокая. Э-э картофель тоже низкий урожайность. Вот, овёс и так далее. И вот мы, значит, видим, как они распределены эти данные. Вот это критерий Хи-квадрат, знаменитый. Мы видим, что содержание гумуса как-то влияет положительно и отрицательно на получение определённых результатов. Мы можем это увидеть в виде различных форм. Ну, в частности, вот есть форма очень наглядная, которую я называю когнитивные функции, которая показывает, какое количество информации содержится в том или ином признаке о том, что объект с этим признаком или под действием этого фактора переходит в определённое состояние. Вот признаки - это содержание гумуса у нас, низкое, среднее, высокое. А это зелёная масса люпина.
Вот мы видим, ребята, цветом здесь изображено количество информации. Значит, э-э красный цвет - это большое количество информации положительное о том, что это наступит событие. Синий цвет, вот такой ярко-синий, фиолетовый, он означает, что это большое количество информации о том, что этого не наступит. Ну, а эти вот это промежуточные. Ну, скажем, там оранжевые такие, зеленоватые, это то, это положительное количество информации о том, что это произойдёт, но меньшее, чем вот здесь.
Значит, вот мы видим, смотрите, что минимальное количество гумуса жёстко обуславливает э-э очень низкую урожайность зелёной массы, получение очень низкой зелёной массы. А вот среднее влияет на получается средняя содержание массы, большое содержание гумуса - большое. Но это в меньшей степени обусловлено, чем низким содержанием э-э гумуса, низкая э-э объём зелёной массы.
И вот мы можем сейчас все эти формы получить, как влияет содержимое, содержание гумуса на все эти результаты. Вот, то есть у нас, кстати, на урожайность картофеля. Оказывается, на картофель очень жёстко влияет большое содержание гумуса. Вот. На овёс - низкое содержание жёстко влияет. На э-э зелёную массу, значит, э-э, на рожь, на рожь тоже большое содержание гумуса жёстко обуславливает большой э-э урожай э-э ржи, значит, и ячмень то же самое.
Вот. То есть вы видите очень наглядные формы, которые показывают, как влияет содержание гумуса на урожайность.

**Раздел 5: Прочие примеры и заключение**

**(Звук уведомления)**

**(Звук уведомления)**

**(Лектор отвечает на звонок)**
Я занят, всё.

**(Лектор продолжает)**

Евгений Вениаминович, по времени у нас уже пара тоже закончилась, как бы. Я заметил, я заметил. 17:05. Ну всё очень интересно рассказали, спасибо большое. Пожалуйста. Если, значит, у вас какие-то есть свои данные, то мы можем такой анализ провести, будут очень интересные результаты получаются. Хорошо. Хорошо. То есть получается, допустим, мы можем узнать, когда зелёная масса люпина большая. А вот тогда, когда вот содержание гумуса максимальное, тогда вот будет максимальное содержание зелёной массы там. То есть мы всё это моментально вытащим, понимаете, всё это узнаем. И этот это же самое, ребята, касается и технологий. Вот у вас там есть какие-то характеристики ваших консервов. Вот эти характеристики от чего-то зависят, от самого процесса технологического. Мы можем узнать, от каких характеристик процесса зависит зависит такие вот результаты, такие, такие.
Ну всё, всего самого хорошего, ребят. До свидания.
Спасибо большое. До свидания. Вам тоже всего хорошего. До свидания. До свидания. До свидания.