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

***Российская Федерация***

***Потемин Илья ПИ2103***

**117 Инженирия знаний и интеллектуалные системы. Лабораторная 5. 2020-10-05**

Обзор Лабораторной Работы №5 (ЭИДОС): Оценка авто с пробегом методом аналогии и требования к зачету

Резюме

Лекция посвящена проведению лабораторной работы №5 по дисциплине "Инженерия знаний и интеллектуальные системы" для группы B 1701 (5 октября 2020 года). Занятие ведут профессора Луценко Е.В. и Аршинов Г.А.

1. Обзор и Подготовка к Лабораторной Работе №5

Предыдущая работа: Подтверждается завершение лабораторной работы 3.02 по интеллектуальной обработке текстов.

Текущая работа (№5): Основная тема – оценка автомобилей с пробегом методом аналогии (риэлторская экспресс-оценка) с использованием системы ЭИДОС. Рассматривается файл исходных данных "Старый Excel", диапазоны классификационных (2-2) и описательных шкал (3-11).

Метод аналогии: Объясняется суть метода: стоимость объекта (автомобиля, недвижимости) определяется путем сравнения с аналогами на рынке с учетом их характеристик (год, пробег, состояние, аварии и т.д.). Сравнивается с затратным подходом (восстановительная стоимость).

2. Работа с Системой ЭИДОС

Установка и Обновление: Проверяется наличие установленной системы ЭИДОС у студентов. Даются инструкции по запуску (start\_aidos.bat или aidos-x.exe) и обязательному обновлению системы до версии от 03.10.2020 через скачивание download.exe с сайта профессора Луценко (lc.kubagro.ru) и разархивирование в папку системы (предварительно выйдя из программы).

Запуск Лабораторной: Необходимо запустить режим 1.11 для очистки предыдущих приложений, затем режим 1.3 для выбора и установки лабораторной работы №5 (или 3.04 по нумерации системы).

Основные Режимы: Упоминаются режимы для просмотра данных (2.1, 2.2, 2.3), создания моделей (5.5) и онлайн-помощи (5.14).

3. Требования к Оценке и Самостоятельная Работа

Получение "5" (Самоэкзамен): Студенты должны самостоятельно разработать облачное ЭИДОС-приложение по своей теме (не из числа существующих лабораторных), создать его описание по шаблону и разместить описание на платформах ResearchGate и РИНЦ.

Другие Оценки: Невыполнение части требований (например, отсутствие описания или публикации) приведет к снижению оценки до "4" или "3". Подчеркивается важность посещения занятий и активной работы для получения положительной оценки. Двойки ставить не планируется, но невыполнение заданий приведет к тройке.

Адаптация Методик: Обсуждается, что созданные методики (формулы) оценки являются эвристическими и требуют адаптации (обновления данных) и локализации (учета региональных особенностей) со временем или при смене условий. Система ЭИДОС позволяет автоматизировать этот процесс.

4. Онлайн-Помощь и Дополнительные Материалы

Режим 5.14: В системе ЭИДОС есть режим онлайн-помощи, где для лабораторной работы №5 (3.04) представлена статья профессора Луценко, подробно описывающая метод риэлторской экспресс-оценки и его применение в системе ЭИДОС на примере оценки автомобилей.

Детальная Расшифровка

1. Введение и Организационные Моменты

1.1. Приветствие и Перекличка

Здравствуйте, ребята. Отвечайте, здравствуйте, здравствуйте. Кто там есть? Здравствуйте. Здравствуйте. Здравствуйте, здравствуйте.

1.2. Дата, Время, Дисциплина и Группа

Сегодня у нас 5 октября 2020 года. Шестая пара, которая идёт с 17:20 до 18:50. На этой шестой паре лабораторная работа номер пять с группой B 1701 по дисциплине Инженерия знаний и интеллектуальные системы.

1.3. Преподаватели

Занятие ведут профессор Луценко Евгений Вениаминович, это я, и профессор Аршинов Георгий Александрович. Георгий Александрович, скажите пару слов студентам. Здравствуйте. Что-нибудь там.

1.4. Посещаемость

Весь курс почему-то не ходит на занятия. Да, что-то как-то маловато. Да. Видно, они все работают допоздна, что ли, или что? Ещё не все просто подключились. Ну, будем ждать, что подключатся. Подключаться надо к началу занятия, а не потом в течение там 15 минут.

2. Обзор и Подготовка к Лабораторной Работе №5

2.1. План Занятия и Обзор Предыдущей Работы

Ну хорошо, ребята. Значит, теперь у нас продолжим занятие по нашему плану. Значит, насколько у меня здесь вот отражено, я отражаю сейчас это прямо в расписании. Вот, что у нас должно быть на этом занятии.

Вот. У меня написано, что мы начинаем лабораторную работу 3.04. Но хотелось бы, чтобы вы подтвердили всё-таки. Вот Марина, ты всё-таки староста, подтверди. Вот мы на последнем занятии изучали работу с текстами, да? Лабораторную работу изучали текст, обработка текстов, да? Интеллектуальная обработка текстов. Так ведь, да? И закончили эту лабораторную работу 3.02. Правильно?

Да. Добрый день, Артём.

Марина. Пожалуйста, ответь. На прошлой лабораторной работе мы изучали… Я отвечала до этого, да, работа с текстами. Вот. Ну мы эту работу прошли. Я вам рассказал про возможности обработки текстов.

2.2. Переход к Новой Лабораторной Работе (№5/3.04 - Оценка Автомобилей)

Теперь у нас осталась ещё одна работа лабораторная. Потом мы будем заниматься тем, что будем выполнять самостоятельную разработку.

Эта работа (3.04/№5) чем она интересна? Тем, что это реальная работа.

2.3. Объяснение Метода Аналогии (Риэлторская Экспресс-Оценка)

Давайте я в двух словах вам расскажу про этот метод риэлторской оценки. Что это такое по методу аналогии?

Вот, допустим, вы пришли покупать машину, вот, на фирму, которая продаёт машины с пробегом. И ходите и смотрите, какие там стоят машины. И там указаны цены на эти машины. Эти цены, они отражают год этой машины, конечно, марку и модель, год, пробег и была ли она в авариях. И также там есть интерес самой фирмы, которая, ну, на этом зарабатывает. Фирма, продающая эти машины с пробегом.

И кто же эти цены назначил, интересно? Их назначили эксперты этой фирмы, специалисты по продаже машин с пробегом. Когда, если бы мы пришли бы не покупать машину, а продавать в эту фирму, и показали бы свою машину им, они бы на что обратили внимание сразу, эти специалисты? Они сразу обратили бы внимание на год её производства, взяли бы техпаспорт, посмотрели бы. Или просто спросили бы, а потом уже до документов когда дойдёт, они посмотрят техпаспорт. Второе, они сразу бы посмотрели, какой у неё пробег. У неё может быть год у этой машины, а пробег 100.000, например. Вот, это, конечно, много для одного года. Ну, мало ли, что там с ней делали. Может, там ездили там в другие города на ней или как такси использовали. Вот. То есть непрерывно она работала, не стояла нигде. А вот весь день она ездила, ездила днём и ночью.

Потом они возьмут такой специальный приборчик, который измеряет толщину краски кузова. И пройдут вокруг этой машины и посмотрят на крылья, на бамперы, на крылья, на капот, на багажник. Везде там подвигают. Даже на крышу могут посмотреть. И найдут сразу места, где машина была повторно покрашена. Даже если это было сделано профессионально, вот, то есть на глаз это ничем не отличается от фирменной покраски. Вот, всё равно они это моментально обнаружат, что там была шпаклёвка, покраска. Вот. И скажут: "Так у вас тут дверь заменена та задняя левая. Или у вас это крыло-то, извините, ребята, с левое спереди, да, оно тоже у вас заменено или бито и потом восстановлено". Вот. Они всё это сразу же вам скажут. Если машина была в аварии, то, в зависимости от серьёзности этой аварии, они постараются определить, откроют капот, посмотрят. Вот, геометрию кузова могут измерить. И сказать вам: "Ну тогда, ребята, вот минус ещё 80 там тысяч, да?" И так далее. То есть вам придут оценку этой машины.

Как они это сделают? Вот сам, сама идея способа оценки какая? Они скажут: такие вот машины, слушайте внимательно, такой марки, такой модели, такого года, с таким пробегом, они на рынке стоят сейчас, вот если вы зайдёте там на Авито или ещё посмотрите другие сайты, вы увидите, что эта машина такая вот, похожая на вашу, стоит столько-то и столько-то. Поэтому вы, ребята, больше не просите. Вот. Ну вы ж понимаете, что мы должны ещё заработать. Поэтому мы вам, если мы такую цену выставим, вот, то вам мы за эту машину заплатим на 50.000 меньше, чем мы будем её, чем та сумма, за которую мы будем её продавать. И вы, значит, соглашаетесь или нет. Вы говорите: "А вот я там в интернете поискал, такие машины стоят за до 400.000, допустим". Они говорят: "Ну тогда идите". Ну они вам предложат, скажем, там 350, скажем, вам на руки за эту машину. А вы говорите: "А вот там я нашёл 450". Они говорят: "Ну тогда, извините, тогда вы не по адресу обратились. Тогда вы берите свою машину, идите и продавайте за 450. А когда не сможете, поймёте, что эта цена завышена, тогда приезжайте к нам ещё раз". Примерно так.

Так вот, что это значит, что такие машины столько стоят? Это значит, что нужно найти аналогичные машины на рынке, именно в этом городе. Это важно. Именно в этом городе. Вот. И посмотреть, сколько они стоят.

То же самое касается, когда мы покупаем недвижимость, например. Вот мы видим, что дом стоит, да, на продажу. И вот его стоимость выставлена. Почему такая выставлена стоимость этого дома? А вот потому, что аналогичные дома в этом месте, вот, в этом году стоят столько вот примерно, как этот дом. Аналогичные по тому, как они сделаны, по какой технологии, из каких материалов, когда. И когда, и в каком именно месте они находятся эти дома. Вот именно вот в этом посёлке там, в этом, на этой, в этом, так сказать, населённом пункте, в этом, в этой части населённого пункта там, на периферии там или в центре. То есть это метод аналогии. То есть мы находим нечто похожее и смотрим, сколько оно стоит. И столько же понимаем, что столько же может стоить и ваш объект.

Какие ещё другие методы оценки существуют? Ещё существует восстановительная стоимость. Сколько стоит такой дом построить? Вот построить такой дом в этом месте стоит дешевле в два раза. Ну, хорошо, пожалуйста. Вы говорите: "А вот почему я должен платить там, допустим, 4 млн? Вот я понимаю, что этот дом стоит 2,5 млн. Вот его построить 2,5 млн. Все расходные материалы, работы, вот всё это вместе столько это будет обойдётся". Нет, хорошо, пожалуйста, понимаем. Берите 2,5 млн, если они у вас есть, и стройте. Покупайте участок и стройте там этот дом. Ради бога, никто вас не заставляет покупать за 4,5 млн. Но вот готовый дом вот здесь вот, уже на участке купленном, вот он стоит 4 млн и всё, понимаете? Вот. Почему? А потому что вот такой рынок, вот это рыночная цена этого дома. А вот эта цена, сколько его стоит построить, это восстановительная цена. Сколько стоит создать такой объект? Да, возможно, это стоит 2,5 млн. А ещё это стоит 2 года, ребята, нервотрёпки и работы, э-э, связанной с созданием, со строительством. Конечно, может быть, вы не сами будете там кладку делать, но вы должны будете найти подрядчика, бригаду какую-то, и за ними поглядывать, как они это всё делают там, начиная с проекта ещё, понимаете? То есть получается, что когда вы платите 4 млн, то за что вы платите? Платите за 2,5 млн за то, что там стоит уже это, за материалы, работу. А ещё за то, что это не 2 года надо будет тратить на это и внимание ещё обращать на это. А прямо сейчас вы его получите. Вот за всё вы это платите. Поэтому восстановительная стоимость, стоимость его создания, она всегда меньше, чем рыночная стоимость. Хотя рынок дело такое тонкое, что если там понастроить столько домов, что они будут стоять целыми улицами пустыми на продажу, тогда может цена упасть даже ниже себестоимости, ниже восстановительной стоимости.

Так что вот такие методы есть: метод аналогий, метод восстановления, восстановительная стоимость есть. Вот, и другие есть подходы. Но, значит, я могу вам сказать, что вот эта задача поиска аналогии, она возникает, правильно? Опытный оценщик, риэлтор, опытный, он посмотрит на ваш этот автомобиль, на ваш дом и сразу моментально, ребята, скажет, сколько это может стоить. Практически, ну, дом там немножко побольше времени займёт, он туда зайдёт вовнутрь, походит, потыкает там стены. Вот. Автомобиль тоже, я сказал, не сразу, а некоторые нужно произвести операции с ним, манипуляции, чтобы понять, сколько он может стоить, то есть исследовать его немножко. Вот. Ну это занимает несколько минут буквально. То же самое касается дома. Может, там только фасад выложен хорошим кирпичом, а внутри там не поймёшь, из чего он сделан там. Это моментально всё выясняется, когда специалист обойдёт этот дом, посмотрит снаружи, изнутри, какие потолки там, какие там перекрытия. Он всё это моментально поймёт. Очень быстро. Поймёт, что старый дом, трухлявый, обложен новым кирпичом итальянским, и его затуляют, будто он сделан вчера по новой технологии там и так далее. Вот, это всё быстро выясняется.

Ну и что дальше можно сказать? Что для проверки, для поиска вот этой аналогии, может быть применён либо естественный интеллект, и опыт эксперта. То есть эксперт просто вот он сам понимает, сколько это может стоить на основе своего опыта, интуиции и профессиональной компетенции. Вот. А можно применить автоматизированную систему.

Ну вот, конечно, абсолютно точных аналогов не существует. Поэтому абсолютно точным аналогом является только сам этот дом, о котором идёт речь. Или, э-э, если дом даже находится рядом, другой, похожий, по похожей технологии сделанный, вот, и по всем параметрам такой же, по объёму помещений, этажности, там, всё-всё-всё очень похоже. Всё равно он находится там 50 м левее и может отличаться там склон, например, уже там начинается участка или ещё что-нибудь. То есть там уже цена может отличаться. Хотя он очень похож и очень место рядом, но всё равно может быть отличие. Поэтому получается вот вопрос в том, насколько этот аналог похож на наш дом, стоимость которого уже известна.

2.4. Специфика Оценки Автомобилей с Пробегом

Теперь, что касается автомобилей с пробегом. Как их оценивают эти автомобили с пробегом? Значит, их оценивают по их характеристикам, которые известны об этих автомобилях. Ну, я уже сказал, марка-модель, год производства, пробег, наличие аварий и тому подобное, отделка салона, рестайлинг, что там с ним делали. Всё это играет роль.

2.5. Разработка Методики Оценки

Ну я могу вам сказать, что разработка такой методики оценки, я вот сейчас взялся в список литературы перешёл этого, этой статьи, прямо сейчас как описание данной лабораторной работы. И здесь я вам показываю, в этой, эта статья посвящена описанию метода, описанию подхода к разработке методики риэлторской оценки объектов недвижимости. И здесь, значит, описывается, как это всё делается. Вот. Для чего я это показываю? Для того, чтобы у вас было представление о том, что вообще-то это не так-то просто делать. Значит, разработка новой методики риэлторской оценки новых объектов – это большая квалифицированная работа. Ну, сейчас я суть этой работы объясню. Допустим, вам фирма риэлторская даёт задание разработать методику риэлторской оценки автомобилей с пробегом. Вы разработчик, который готов разработать такую методику для этой фирмы и имеете соответствующую квалификацию и готовы взяться за эту работу и выполнить. И вы делаете что? Вы начинаете с того, что вы берёте, разрабатываете формулу, которую можно использовать для оценки этого автомобиля, к примеру, там или квартиры, смотря, для чего предназначена эта методика. Ну, допустим, это будет методика оценки стоимости автомобиля.

Как обычно к этому подходят вопросу? Ну, примерно так. Вот у этого автомобиля есть базовая цена. Причём базовая цена не в том году, когда его купили, а вот сейчас. Вот если сейчас пойти в фирму, где вы купили этот автомобиль, и там купить аналогичный автомобиль новый сейчас. Вот. Ну я так почему говорю, я вам могу сказать, что, скажем, я 3 года назад купил машину Hyundai Solaris в салоне, в автосалоне там на Аэропортовской за 820.000. А вчера я там был в этой, в этом салоне и видел такую же точно машину, абсолютно один к одному, с таким же двигателем, с такой же коробкой, все параметры совпадают, с таким же салоном, ребята, с таким же электронным оборудованием, с таким же этим мониторчиком там на приборном щитке. В общем, абсолютно всё один к одному. А она стоит 1.020.000. Я спрашиваю сотрудников этого, менеджеров этого автосалона, спрашиваю: "А почему? Я может, в чём-то ошибаюсь? Вот эта машина, она ведь стоила 3 года назад, в семнадцатом году, 820.000". Они говорят: "Да, вы не ошибаетесь. Она стоила 820.000". "А почему сейчас она стоит 1.020.000?" "А вы вот что, не в курсе, что ли, что с евро происходит? У нас же все комплектующие, вся поставка полностью из-за рубежа идёт. 100%. Это же зарубежная машина. Машина зарубежного производства. Ну, если не считать, что она делается в Петербурге. Ну, будем считать, что действительно она зарубежного производства. Допустим, это Kia, а не Hyundai". Вот. И они говорят: "А вы знаете, что когда вы покупали тогда, в семнадцатом году, евро стоил 68 руб., а сейчас он стоит 90 руб.?" То есть она вообще-то должна быть ещё дороже. Это мы ещё, так сказать, немножко не настолько повысили её цену, как она должна была бы быть повышена, если бы просто курс рубля мы учитывали и всё. Мы ещё учитывали другие обстоятельства и постарались повысить её не так сильно, как бы она чисто вот по курсу получилась бы.

Вот. Ну, то есть получается так, что мы должны взять базовую цену сейчас в этом на такую машину. Вот. Потом, ну если мы берём ту цену, которая была в автосалоне, вот, тогда мы должны, то есть когда мы можем взять ту цену, которая была в автосалоне 3 года назад? Только в том случае, если сейчас она не производится эта машина, не продаётся новая. Новая машина не продаётся. Вот. То есть мы берём машину, стоимость машины. Ну, допустим, даже 3 года назад. А потом берём и добавляем какие-то слагаемые с разными знаками, с плюсами и с минусами к этой базовой цене. Значит, если каждый год, прошло несколько лет, то каждый год уменьшает стоимость машины на 10%. Вот. Потом просто машина выехала за ограду салона. То есть её купили и выехали из автосалона. И потом подумали: "Ой, зря мы её купили. Давай-ка мы её обратно отдадим". Вот. И, значит, если вы её сразу же попытаетесь продать, она будет стоить на 10% дешевле та машина будет. Просто потому, что вы выехали и заехали. Вот. То есть если вы попытаетесь её сдать вот в этот, э-э, в эту фирму, которая продаёт машины с пробегом, они никогда её не возьмут по той цене, какая стоит новая машина, потому что она стоит как новая вот через дорогу в самом автосалоне. А в этой фирме она стоит дешевле, потому что это машина с пробегом.

Вот. Теперь, э-э, каждый год уменьшает цену машины на 10%, допустим. Потом смотрим, какие характеристики этой машины, пробег какой. Каждые там 1.000 км тоже уменьшает стоимость вот настолько-то. А если она была в ДТП, то в зависимости от тяжести ДТП ещё уменьшается её стоимость. А если, значит, мы смотрим, какое у неё, бывают разные комплектации у этих машин. Какая коробка, какая отделка салона. Смотрим на это всё. И говорим: "Так, а вот с такой отделкой салона, тогда это вот дешевле. А если с такой, тогда дороже". Вот. И смотрим на был ли у неё рестайлинг. То есть что-то там с ней делали, улучшали её внешний вид там специально как-то. Бывает иногда какие-то работы проводят кузовные, чтобы улучшить машину, сделать её более модной, например, там или как-то более спортивный вид ей придать. Тоже это влияет на стоимость, увеличение стоимости. А что-нибудь с салоном делали? Покупали там чехлы там или, может быть, другие какие-то, какая-то другая отделка там специальная была осуществлена. Вот. Если там кожаный салон, например, там или ещё что-то там, то тоже это влияет на цену.

Короче говоря, мы должны вывести формулу, добавляя к базовой цене и отнимая от неё различные слагаемые, отражающие динамику цены. Потом эту формулу, значит, мы должны проверить. Здесь вот как раз эта методика описана, но я не буду ссылку вам бросать, потому что она есть в списке литературы. Получается некая формула, которая оценивает стоимость этого объекта движимости или недвижимости. И если мы эту формулу начинаем проверять на различных объектах аналогичных, то она должна выдавать разумные значения. Такая методика, вообще-то, разработка такой методики – это дело профессионалов, риэлторов, имеющих опыт, и даже не просто риэлторов, а именно вот разработчиков. И, в общем-то, является достаточно квалифицированной работой и может стоить деньги хорошие. Вот. Ну, самые простецкие такие методики, они могут быть там, допустим, разработаны за месяц-два, например, работы. Именно вот методики риэлторской оценки. Вот. И стоит, ну, скажем, там 300-400.000 руб. Это как минимум, так сказать. Наверное. Ну, хотя я, честно сказать, с рынком таких разработки таких методик не знаком сейчас, но вообще-то вот стоит это денег эта работа.

И вот возникает такая ситуация. Мы спрашиваем, а вы когда разрабатываете эту методику, разработчиков самих спрашиваем. Когда вы разрабатываете эту методику, вы применяете какие-то автоматизированные системы для вывода вот этих вот формул? Для того, чтобы оценить, как влияют те или иные свойства этих объектов на их цену на рынке сейчас. Нет, нет, мы просто вот понимаем, что как-то это влияет, и пытаемся подобрать, подогнать, проверяем это. У нас формула выдаёт результат с некоторой ошибкой. Э-э, постоянно в сторону увеличения там или уменьшения. Мы смотрим, что какие-то необходимо ввести поправки в коэффициенты этих формул. Ну, то есть, если так сказать по-простому, то это делается с помощью подгонки. Это самая такая, самая культурная формулировка. То есть эта формула подгоняется, то есть она имеет эвристический вид. То есть теоретически она не обоснована никак. То есть это эвристические формулы, то есть они как бы, скажем, вводятся на основе интуиции разработчиков, опыта и интуиции разработчиков, которые понимают, что вот эти вот такие-то свойства удорожают, такие-то свойства удешевляют объект, который будет оцениваться. И э-э, они выводят эту формулу, а потом её подгоняют, её уточняют её. И когда она начинает работать уже с приемлемой погрешностью, уже практически точно даёт оценку, тогда её сдают заказчику. И заказчик начинает ей пользоваться.

3. Требования к Оценке и Самостоятельная Работа

3.1. Условия Получения Оценок

Я ещё раз хочу напомнить вам, ребята, что те, кто выполнят самостоятельную разработку и поместят её описание в ResearchGate и в РИНЦ, получат самоэкзамен, пятёрку. А те, кто этого не сделают, пятёрку не получат. Это у нас и есть то самое задание, которое нужно выполнить, чтобы получить пятёрку. Причём эта работа не должна здесь вот быть среди уже имеющихся работ. Должна быть новая работа.

Вот ссылочка на инструкцию по тому, как мы будем это делать, эти работы. Добрый вечер. Ссылки я вам давал. Сообщаю, что сейчас уточнилась немножко информация по поводу того, как мы будем принимать что там у нас, экзамен или что там. Вот. То есть я вам повторяю, что те, кто сами, ну, вернее так, не то что сами, а под нашим руководством создадут облачное ЭИДОС-приложение в соответствии с этими инструкциями и с нашей помощью. Вот, но всё-таки самостоятельно это будут делать, не мы будем за них делать. И разместят описание, сделают описание по шаблону этих приложений и разместят его в облаке, э-э, в ResearchGate и в ЭИДОС, э-э, в облаке ЭИДОС-приложений, облачных ЭИДОС-приложений. А также это описание разместят в ResearchGate и в РИНЦе, получат самоэкзамен. Если чего-то этого не будет из этого хватать, ну, допустим, приложение сделают, а описание нет, пятёрка не будет уже поставлена. Будет там четвёрка. Если, если, допустим, приложение готово, а описание не готово. А если, значит, ещё и приложения нет, ну тогда тройка.

3.2. Важность Посещения и Активности

Вот. Если вообще не будете ходить на занятия, как вот сейчас мы это видим, то можете передать, что те, кто вот не ходят или не сделает чего-то из того, что я сейчас сказал, всё равно получат тройку. Двойку я не поставлю. Вот. Ну и четвёрку, пятёрку не получат. Вот такие вот дела.

4. Работа с Системой ЭИДОС

4.1. Запуск и Проверка Установки

Теперь начинаем саму работу. Саму работу. Для этого переходим в папочку системы. У вас у всех система должна быть установлена на компьютерах. Сейчас в чате, пожалуйста, поставьте минусики те, у кого она не установлена в чате. Уже должна быть у всех она установлена. Запускаем старт ЭИДОС, файлик.

4.2. Обновление Системы (при необходимости)

Если у нас это не получается сделать, какие-то возникают ошибки, тогда запускаем исполнимый модуль системы aidos-x.exe. Но тогда на моём сайте вот здесь, lc.kubagro.ru, второй пункт, скачать и запустить систему ЭИДОС. Вот тогда сначала вам нужно скачать обновление системы, файл download.exe. Скачиваете обновление, разархивируете эти обновления в папочке системы с исполнимым модулем. При этом нужно выйти из самой системы, потому что иначе она не будет обновлена. Чтобы обновить исполнимый модуль системы, надо, чтобы он был не был запущенным. Был не запущенным этот модуль. Обновляете его. Вот. Этот исполнимый модуль и другие программы. Там есть несколько модулей внешних, а также файлы, которые система отображает. В общем, вот всё полностью разворачиваете файл download.exe.

4.3. Проверка Версии Системы

У вас должно быть вот здесь вот вверху написано: дата бета-версии 03.10.2020. То есть система должна быть 03.10.2020.

4.4. Запуск Лабораторной Работы №5 (3.04)

Потом заходим в режим 1.3. Я сейчас посмотрю, у кого у вас система не стоит. Чат. То есть у вас у всех стоит, раз минусов здесь нет, да? Сначала запускаем режим 1.11, стираем все приложения, какие там были. Потом заходим в режим 1.3 и выбираем работу 3.04.

4.5. Параметры Лабораторной Работы

Сразу же обратите внимание на параметры какие. Файл исходных данных Старый Excel. Диапазон классификационных шкал со второй по вторую. И диапазон описательных шкал с третьей по одиннадцатую. Все остальные параметры такие в упрощённом, упрощённые варианты этих параметров вы видите.

И мы видим здесь впервые, может быть, что у нас есть числовые шкалы и классификационные, и описательные. В числовых шкалах мы берём, задаём три классификационных шкалы. То есть три три градации числовых, три числовых градации в классификационных шкалах. И в описательных шкалах задаём 12 числовых градаций, диапазонов. При этом у нас получается матрица моделей: три колоночки в матрице моделей, три класса. Вот. И 272 э-э, 272 строчки. То есть это значение свойств или значение факторов. То есть это довольно много уже для такой числовой модели.

4.6. Процесс Создания Модели

Выходим на создание модели. Ребят, если у вас есть вопросы, пожалуйста, можете спрашивать, потому что лабораторная работа. Вы имеете полное право сейчас задавать вопросы по ходу моего изложения. Если вам что-то непонятно, спрашивайте. Вы сейчас на своих компьютерах должны это всё повторять, чтобы научиться не смотреть на экран, а научиться именно всё это делать.

Что произошло? Я сейчас так рассказываю очень упрощённо, сокращённо, потому что первая работа у нас была посвящена подробнейшему изучению всех этих вопросов. Я всё очень подробно объяснял. А сейчас мы идём достаточно быстро всё. Но у нас будет и новое по этой лабораторной работе, которое сейчас я расскажу вам дальше. А пока что мы вот делаем то, что уже должно быть у вас навыки должны уже возникнуть, отработано всё.

Вот. Значит, что сейчас было сделано автоматизировано, в автоматизированном режиме? Сформированы классификационные описательные шкалы и градации на основе файла исходных данных. И с их использованием закодированы исходные данные, сформирована база событий и обучающая выборка. А потом всё это проиндексировано.

4.7. Дальнейшие Шаги по Инструкции

Теперь видите, у нас есть сообщение студентам о том, что делать дальше. Сейчас мы это и будем делать, то, что здесь нам рекомендуется делать дальше. А дальше нам рекомендуется что, что делать? Дальше нам рекомендуется посмотреть на файл исходных данных, посмотреть сами в режимах 2.1, 2.2, 2.3 и так далее, посмотреть классификационные описательные шкалы и градации, обучающую выборку, создать модели и так далее, и так далее, то, что мы уже делали, но только на этих новых данных.

Значит, да, и у нас появилось здесь в диспетчере приложений сообщение, что у нас установлена лабораторная работа 3.04, оценка автомобилей с пробегом по их характеристикам. Вот. Теперь смотрим то, что нам было рекомендовано, делаем. То есть открываем файл исходных данных. Файл такой своеобразный.

5. Онлайн-Помощь и Дополнительные Материалы

5.1. Режим Онлайн-Помощи (5.14)

Значит, и, значит, я вам сейчас скажу вот ещё что. Обратите внимание, что у нас есть режим в системе 5.14. Это онлайн-хелпер по лабораторным работам. Онлайн-хелпер. И в этом режиме мы видим все лабораторные работы перечислены всех трёх типов. И есть лабораторная работа 3.04. И по ней мы открываем онлайн-хелп.

5.2. Статья Профессора Луценко

В качестве этого онлайн-хелпа используется статья, которую я написал 7 лет назад. И эта статья называется "Разработка без программирования и применение в адаптивном режиме методик риэлторской экспресс-оценки по методу аналогий, метод сравнительных продаж, в системно-когнитивном анализе и интеллектуальной системе ЭИДОС". И это сделано на основе примера автомобилей с пробегом.

6. Практическая Работа с Данными и Моделями

6.1. Просмотр Исходных Данных и Шкал

Значит, мы видим, что у нас три ценовых диапазона. Ну, что я могу сказать? Что, конечно, они эти диапазоны очень велики. И можно лишь сказать, что это низкая цена, средняя или высокая. Потом я прокомментирую этот момент ещё раз. Теперь смотрите, марка машины, фирма, грубо говоря. Это уже и марка, и модель какая. Кузов какой. Пробег и так далее. Значит, если мы увидим год, год выпуска как число воспринимается, и здесь уже диапазон, диапазон в годах указан. Пробег то же самое, это числовая колонка от вообще отсутствия пробега практически, 1800 км – это вообще не пробег, это пробег новой машины практически. И такой приличный пробег – 460.000 уже. Объём двигателя от единички до 5,7. Ну, цвета у них самые разные, коробки вот тоже разные. И отделка салонов тоже разная: велюр, кожа, комбинированная, ткань.

6.2. Просмотр Обучающей Выборки и Моделей

Смотрим обучающую выборку. Вот у нас вверху, в верхнем окошке название объекта обучающей выборки. А в левом нижнем окошке класс ценовой диапазон, к которому он относится. А справа мы видим в окошке коды свойств, то есть значение свойств этого автомобиля. Вот. И видим также другой вариант этой обучающей выборки, который я вам показывал, что он может вот так выглядеть. И здесь у нас уже побольше, в общем-то, строчек.

Дальше мы осуществляем синтез и верификацию модели. Ребят, я хотел вам сейчас сообщить такую новость, что я вот в этой новой версии исправил эту неточность в расчёте модели ПРЦ-2, которая возникала при расчёте на графическом процессоре в предыдущих версиях. Хотя эта неточность не сказывалась на достоверности модели, вот, но было немножко неприятно. Я вас предупреждал об этом. Вот. А сейчас можно спокойно запускать на графическом процессоре.

Вот так вот считаются модели сейчас все и проверяются на достоверность. Это занимает немножко побольше времени, чем в тех задачах, которые мы изучали на предыдущих лабораторных, потому что эта задача, в общем-то, относится уже не к учебным, но она такая, на грани, что-то можем на стыке научной, то есть задачи и учебной. И причём научная задача, которая вполне реальна и имеет практическое значение. 43 секунды отошло. То есть у нас созданы сейчас модели, которые мы знаем, как их посмотреть в режиме 5.5. Мы эти модели видим. Вот. Видим, каких машин, какого ценового диапазона, сколько представлено. У нас 366 машин низкого ценового диапазона, значит, 32 – среднего ценового диапазона и три – высокого ценового диапазона машины. И мы видим, какое количество информации содержится в признаках этих машин, о принадлежности или непринадлежности их к каждому из классов. Здесь есть и с минусом, видите, много. Вот. И видим разные модели. Это была инф-1, это модель Хи-квадрат. У ней такая особенность у этой модели Хи-квадрат, что сумма и средняя равна нулю. Но они, если они точно равны нулю, тогда они не выводятся вот здесь в форме. А здесь они равны нулю почти что, скажем так. Возможно, там в седьмом знаке они не равны нулю, поэтому отображаются вот в такой форме это. То есть это не точно равно нулю, а почти равно нулю значение. И в тех знаках, которые отображаются, мы не видим там неравного нулю, но где-то там есть дальше. Вот. Ну, в общем, так вот примерно. То есть мы видим эти модели.

6.3. Оценка Достоверности Моделей

Теперь посмотрим на достоверность этих моделей. То есть мы действуем по тому, идём по тому пути, который мы уже несколько раз, два раза уже, точнее, мы уже по нему проходили по этому пути. Вот. Хотим узнать, как влияют свойства машин с пробегом на их ценовую категорию. И вот те свойства, которые я вам перечислял. Для этого мы провели формализацию предметной области, то есть исходные данные собрали. Потом ввели в систему и с помощью автоматизированного программного интерфейса. При этом были созданы справочники классификационных описательных шкал и градаций. С их помощью закодирована обучающая выборка. Потом были созданы модели статистические три модели, системно-когнитивные. И сейчас мы посмотрим, какая их достоверность. Потом выберем наиболее достоверную, будем решать задачи идентификации ценовой категории и исследования предметной области, куда входят и решение обратной задачи прогнозирования. Ну, в общем, короче, будем смотреть, какие классы сформировались, попросту если сказать.

Вот мы смотрим, какая у нас достоверность модели. Значит, видим, что модели получились не чрезвычайно, не очень высокой достоверности. Значит, ребят, я могу вам сказать, что когда я эту задачу решал, ну, у меня не было такого стремления создать очень хорошую методику прогнозирования или оценки ценовой категории автомобиля с пробегом. Если бы я такую задачу ставил, то я бы ещё бы потом провёл бы ряд работ. Я бы, знаете, что сделал? Я бы сузил эту задачу, то есть не только, то есть не включал бы в неё машины разных марок, а сделал бы эту методику для машин одной марки, ну, например, машины Toyota, к примеру, или машины Hyundai, понимаете? Потому что, сейчас вы увидите, какие там ошмётки возникают, и согласитесь с тем, что я был прав сейчас, когда это говорил.

Значит, что мы здесь видим? Смотрите, ребята. Это, значит, у нас графики, которые отражают количество истинно положительных, истинно отрицательных решений, а также ложноположительных и ложноотрицательных при разных уровнях сходства. Мы видим, что картина довольно сложная здесь получилась. Значит, вот до сначала положительные решения с положительным уровнем сходства. До 20% примерно уровня сходства преобладают ложные решения. Истинные есть решения, но их меньше, чем ложных. Начиная с 20 там 2% примерно, истинных решений больше, чем ложных. Но ложные наблюдаются вплоть до 95%. То есть они наблюдаются постоянно. То есть они есть, но их хоть и меньше, чем истинных, но они есть. Что касается решений о непринадлежности автомобиля к той или иной ценовой категории, здесь картина более такая приятная. Здесь всегда практически больше истинных решений. Ну и ложные тоже есть. Причём иногда при больших уровнях различия от 75 там до 65% различия, примерно одинаково истинных и ложных решений. В остальных уровнях сходства, ну ещё здесь вот где-то где 18 тоже примерно одинаково. Мы можем это увидеть в таком виде вот, где уже разность показана. Видите, при увеличении уровня сходства практически закономерно растёт доля истинных решений. Но не совсем линейно она растёт. Даже не монотонно, то есть даже бывают уровни сходства, при которых понижается опять разница между истинными и ложными решениями. И даже вот здесь вот при решениях отрицательных, даже она может даже может быть больше становиться ложных решений. Есть небольшой там кусочек, когда вот ниже опустились. Правда, здесь мы можем задать интервал сглаживания побольше. Тогда, возможно, картина будет более такая симпатичная. Ну, так вот, в общем, даёт, даёт нам представление о том, как у нас распределены эти решения. Ну и возьмём модель вот эту, вроде здесь по критерию Ван Ритбергена получше. N4. Ну я вот N3 всё-таки модель советую использовать, потому что очень, ну более логично, более наглядно. Мне кажется, что и более правильно она работает. То, что я вам показывал, только с большим, большей величиной уровня интервала сглаживания.

6.4. Идентификация Ценовой Категории и Исследование Предметной Области

Вот. И, значит, дальше мы что должны сделать, чтобы решить задачу идентификации ценовой категории? Мы должны выбрать текущую модель. В качестве которой мы выбираем модель F3. Вот. И потом решить задачу идентификации в этой модели выбранной. Потом мы можем смотреть на результат решения задачи идентификации ценовой категории, то есть оценки автомобилей.

И дальше вот смотрите. Значит, здесь вот у нас идут слева объекты обучающей выборки, которые были скопированы в распознаваемую и определена идентификация ценовой категории. И мы видим, что, в общем-то, есть и истинные решения, и ложные. Ну как вот мы видели из графиков этих. Вот. Ценовая категория дешёвые машины. И мы здесь видим, что довольно много машин правильно идентифицировано с этой категорией. А к ней в основном и относятся все машины этой обучающей выборки. А средняя ценовая категория? Тоже так более-менее, да? А высокая ценовая категория? А здесь у нас много ошибок. Видите, тут у нас две машины правильно отнесены, остальные не совсем, мягко сказать. То есть с высокой ценовой категорией проблемы у нас.

Есть форма интересная – достоверность идентификации по классам. Вот мы можем взять, рассортировать по F-мере. И тогда мы видим, что у нас дешёвые машины очень неплохо идентифицируются. Средняя ценовая категория чуть похуже. И есть средневзвешенная достоверность идентификации машин всех категорий. А вот высокая ценовая категория, она идентифицируется хуже, чем средневзвешенная. Ну, то есть можно сказать так: дешёвые и средние более-менее правильно оцениваются, а с высокой ценовой категорией больше ошибок, и в основном они связаны именно с этой категорией.

6.5. Анализ Влияния Свойств (Обратная Задача/Кластеры)

Что же это значит, ребята? Почему так интересно у нас получилось? Значит, смотрим на решение обратной задачи прогнозирования. Но здесь это задача при статичной интерпретации. Здесь же у нас не факторы мы рассматриваем и влияние этих факторов на будущее состояние объекта моделирования. Мы же рассматриваем значение свойств объекта моделирования. И какое количество информации в этих значениях свойств содержится о принадлежности этого объекта к тем или иным классам. Вот. И вот мы видим, значит, то есть это обратная задача, то есть это просто вывод информации о том, какие классы сформированы, вот так скажем.

Значит, смотрите, берём мы низкую ценовую категорию. Берём модель F3. И смотрим, что наиболее характерно слева для этой ценовой категории и что наиболее нехарактерно. Ну и в графическом виде это смотрим. Это вот решение, ребята, решение той задачи, которую риэлторы решают, разработчики риэлторских оценок решают с помощью своего естественного интеллекта. Они должны понять, какие свойства автомобиля, как влияют на ценовую категорию.

Вот мы видим, ценовая категория низкая, ну то есть машины дешёвые. Ну, тут действительно довольно дешёвые, и, скажем так, не совсем дешёвые, но всё-таки из тех вот цен, которые там были в модели, можно считать, что они дешевле остальных. Что мы видим характерно для этих машин? Очень характерно, что отделка салона ткань, кузов седан, маленький объём двигателя 1.4, 1.8. Ну, не минимальный, но довольно низкий. Механика коробка. Вот. Объём двигателя чуть побольше может быть, 1.8, 2.2. Вот. Это машины старые, ребята. То есть у них из двенадцати градаций по году выпуска – это восьмая градация, то есть довольно старые машины. 2006 года, 2007. И машины с довольно приличным пробегом от 78.000 до 116. А что для них нехарактерно для такой ценовой категории? Отделка салон кожа, кузов джип, автомат коробка-автомат. И нехарактерно, чтобы это были машины Toyota Prado. То есть они не дешёвые в ценовой категории. Нехарактерно для этой же категории дешёвых машин очень большой объём двигателя – 4.5, 4.9. Ну ясно, что это характерно для джипов как раз. Или просто большой. Это очень большой, 10 из 12, да, последняя почти что интервальное значение. Или просто большой – 3.4 л, 3.7 л. И это нехарактерно, ребят, для новых машин. Ну, сравнительно новых машин. Ну это, конечно, не новые машины, но на тот, это сейчас не новые. Но на тот год, когда статья была написана, 7 лет назад я делал эту работу, вот эти машины, они были не очень старые. 2008, 2010 года. Но здесь было бы правильнее написать не год выпуска, а возраст машины, наверное, было бы более правильно. То есть нам надо было бы из текущего года вычесть год выпуска. Ну я тогда этого не сделал.

То есть это о чём говорит? О том, что мы можем эту методику совершенствовать. Если мы её сделаем отдельно для машин разных категорий ценовых, вернее, не ценовых категорий, отдельных разных моделей, то тогда мы увидим, что она будет гораздо более достоверна.

Теперь смотрим среднюю ценовую категорию. Для неё что характерно? А здесь уже как раз вот мы видим, что и кожаный салон, и джип, и автомат-коробка, и средний объём двигателя. Не маленький, а 3.4 – это не маленький объём, средний. Это седьмая градация из двенадцати. Либо большой объём двигателя – 4.5, 4.9. И машины не старые.

И смотрим, а что же у нас характерно для самой высокой ценовой категории, самых дорогих машин? Мы видим, что это, значит, цвет сюда попался белый. Это довольно новые машины для того года, когда статья была написана. С машины с небольшим пробегом. Опять же, новые машины. С коробкой автомат. И, видите, сразу модель выскочила BMW X6 и Porsche Panamera, понимаете? То есть это машины, которые высокой ценовой категории. То есть это выше 3 млн стоимость, до 4,5 млн. Как квартира, в принципе. Или дом.

Вот. Теперь смотрим, какой смысл тех или иных признаков, этих машин. То есть значение свойств автомобилей. Как влияет цвет автомобиля на стоимость? Интересно посмотреть. Так, кстати, мы можем это посмотреть путём фильтрации на предыдущем, в предыдущем режиме свода анализа. Берём вот, допустим, э-э, как влияет на принадлежность к дешёвой ценовой категории цвет машины? Вот. Цвет. И видим, что, ну, видим, во-первых, что он не очень сильно влияет цвет. Всё-таки для дешёвых машин более характерен серебристый цвет, серый, а также вот красный, синий, светло-голубой. А вот для дешёвых нехарактерным чёрным является и белый. То есть какие цвета более характерны, какие менее характерны. Кстати, ребят, я когда-то провёл… Также можно любое совершенно свойство изучать, как влияет там год выпуска, например, на стоимость. Что на низкую ценовую категорию больше влияет высокое значение года выпуска. Вот, то есть это старые машины. Вот. И нехарактерно десятое и одиннадцатое диапазон года. То есть это новые машины. То есть новые машины, короче говоря, стоят дороже, они не относятся к низкой ценовой категории. А старые машины относятся к низкой ценовой категории.

Вот. Теперь смотрим Porsche, например, да, Porsche. Это у нас либо очень дорогая машина, либо средняя по цене. И уж точно не дешёвая. Вот. Видите? То есть я о чём хочу сказать? Что система, несмотря на то, что мы дали такой вот ей материал сыроватый, я бы сказал, по очень многим маркам и моделям, вот по тем характеристикам, которые вы видели, она даёт вполне такие разумные результаты исследования. То есть похоже, как если бы вот эксперт сидел здесь, риэлтор этой фирмы, которая там машины берёт по обмену. Обычно они к тому же самому салону относятся, который, ну или к нескольким салонам, которые рядом продают новые машины, а эта фирма с ними сотрудничает, продающая машины с пробегом. Вот покупатель приезжает на своей машине, туда свою машину сдаёт старую, а берёт новую. И засчитывается это ему в стоимость. Вот обычно они вот по этой схеме работают. И вот если это показать эксперту, оценщику в этой фирме, которая берёт машины с пробегом и продаёт их, он скажет, это, в принципе, то, что я и так знаю. Вот это всё, что сейчас вот я вам показывал, это вполне разумные вещи, которые оценщик этой фирмы и сам знает, как эксперт.

7. Адаптация и Локализация Методик в ЭИДОС

Но я вам могу сказать, что, во-первых, это сделано на основе конкретного прайс-листа конкретной фирмы в конкретном году. И эти вот формы все, они ведь не просто рисуются на качественном уровне графическом. Мы видим, что вот это вот так знает вот это, что это вот точно не дешёвая, то есть Porsche – это точно вот не относится к ценовой категории к низкой, а либо к, скорее всего, к высокой. Ну, может быть, и к средней. Вот. А мы же мы видели здесь вот коэффициенты числовые. Ведь у нас же не просто она даёт на качественном уровне оценку этого, а она ж даёт это и количественную оценку для всех этих значений свойств, какое количество информации в них содержится о тех или иных моментах. И вот эксперт, он не даёт количественные оценки, понимаете? Он даёт на качественном уровне. Ну потом он, конечно, выражает это в числе каком-то, стоимость этой машины. Но как влияет каждое из этих свойств, он, в общем-то, понимает на таком качественном уровне.

И вот разработчики риэлторских методик риэлторской оценки, типа той, что я вот показал вам в статье, они тоже, в общем, на качественном уровне знают это всё. И свои эти формулы выводят, да, которые имеют пригодность в данном месте, в данное время. А когда их необходимо модифицировать, опять надо их звать, что ли?

А здесь, если мы используем автоматизированную систему, я хочу подчеркнуть, ребята, очень важную мысль. Вот сейчас, пожалуйста, послушайте внимательно. Система ЭИДОС является не просто системой, где мы эксплуатируем какую-то методику, которая создана. Но мы и эксплуатируем её в той среде, в которой она создана. То есть это инструмент создания таких методик и инструмент их, естественно, модификации. То есть можно их адаптировать и локализовать. Что такое адаптировать? Это учесть динамику предметной области. То есть через какое-то время эту методику нужно пересчитать модели. Сейчас я объясню, как это делается. А что такое локализовать? Вот взяли, повезли из Краснодара, повезли там, ну, скажем, в Кореновск, например. И там тоже есть фирма, торгующая машинами. А там цены немножко выше, чем в Краснодаре. В той фирме. Почему? Ну, потому что они берут машины из Краснодара просто. Если это попросту сказать, они берут машины из Краснодара, везут их в Кореновск, и там ими торгуют. То есть они их покупают здесь, ну примерно как покупатели вот покупают машины, люди, физические лица. Так же вот примерно они приходят, покупают у себя, их продают. И зарабатывают за счёт того, что там немножко дороже они, эти машины стоят. А если поехать в Москву или в Петербург, или во Владивосток, или в Хабаровск, то там ещё дешевле машины стоят. Вот на Дальнем Востоке будут очень недорого стоить китайские и японские машины, дешевле, чем у нас заметно они будут стоить. Потому что пока сюда довезёшь, они дорожают. То же самое касается Москвы, Петербурга, где рынок перенасыщен, и там цены ниже, соответственно. Ну, то есть спрос там, конечно, огромный спрос, огромное население там, причём с деньгами. Но у них, так сказать, девать эти машины девать некуда там, понимаете? Там в эти салоны забиты этими машинами. И они очень часто чем занимаются? Тем, что просто эти машины продают на периферию, ну, допустим, в Краснодар. А здесь их перепродают. Но при этом, то есть откуда берутся машины, которые продаются в Краснодаре? Кто-нибудь задумывался, откуда их привозят? Очень часто их привозят, ну, может, конечно, на корабле привезут откуда-нибудь, но, в принципе, их очень часто из Москвы, Петербурга и привозят. Причём в Петербурге их там и заводы уже есть, где Nissan там построили заводы, и Hyundai построили. Вот. Просто там производят эти машины. И везут сюда. Причём производят в основном из комплектующих самих этих поставщиков этой владельцев этой торговой марки. То есть марки Nissan или марки Hyundai. Иногда бывает там другой какой-нибудь ставят шильдик. Вот. И потом говорят, что это на конвейере Nissan производят там и так далее.

Так вот, э-э, надо её, если эту методику применять где-то в другом городе, то надо её локализовать на данных из фирмы, которая торгует в этом городе. Если проходит какое-то время, то надо её адаптировать. Как осуществляется адаптация и локализация? А очень просто. Просто вводятся данные этого места, этого времени в систему. И методика пересоздаётся. При этом вопрос возникает такой: а будет ли она работать сразу? Взять вот сейчас вот эту методику и на ней начать эксплуатацию. Будет работать она и сразу, но с погрешностью. И, возможно, не стоит ей пользоваться из-за того, что эта погрешность может быть велика. Но если мы будем удалять старые данные, а новые добавлять в базу исходных данных, на основе которых эти модели созданы, то постепенно эта модель локализуется и адаптируется. И потом дальше может быть использована в адаптивном режиме. То есть мы взяли, то есть регулярно, ну, может быть, раз в неделю, например, там в субботу, воскресенье, к примеру, вот, можно просто брать и пересоздавать модели на, добавив туда новые данные в эти исходные данные, а старые удалив. И остановиться там на каком-то количестве вот этих данных, там 500, например, там или 1.000 примеров машин. И всё, и придерживаться 1.000. Старые удаляем, новые добавляем. И добавляем те цены, по которым реально вот их продаются эти машины. Что поставишь выше цену, она будет там стоять до скончания века. Поставишь ниже, её мгновенно заберут и скажут, какие-то странные продавцы вообще там за так продают машины. Наверняка там что-нибудь не то. Возникает вопрос: а что ж там не то? А может, там всё то, просто они взяли, ошиблись в цене и продали дешевле, чем она стоит машина. Примерно так вот.

8. Заключение и Вопросы

Теперь смотрим на какие могут быть, какое может быть исследование проведено этой модели. Вот, ну то, как влияет на ценовую категорию те или иные факторы, мы вот можем посмотреть. Здесь факторов много, поэтому, я думаю, значения свойств, поэтому видно, наверное, не будет ничего. Вот. Это информация о том, что надо вот здесь вот поставить, ну хотя бы 48 там вот так вот. Ещё тогда можно понять, что там написано. Хотя тоже уже не особо. Ну а лучше всего 32. Вот это тоже из опыта знаю. Вот 32 это всё читабельно. Вот тут те значения свойств, которые наиболее характерны, наиболее нехарактерны для той или иной ценовой категории. Вот ткань характерна для дешёвых машин, а кожа нет. Джип нехарактерно и коробка-автомат. А, значит, седан и маленький объём двигателя и механика характерна. Ну, в принципе, соответствует интуитивным представлениям. Вот. И можно сеть посмотреть, слой нейронной сети. Как она выглядит. Здесь мы увидим, как влияют свойства машин на принадлежность, непринадлежность ценовой категории. Три ценовых категории, и мы видим, какие значения свойств, как влияют. Для этой нехарактерно, для этой характерно. Мы видим, что высокая ценовая категория слабо детерминирована. То есть низкая и средняя более-менее описаны вот этими характеристиками, а высокая – нет сильных связей, видите, по модулю. То есть есть связи по модулю не велики. Это признак того, что эта категория плохо описана, плохо идентифицируется. Нет наиболее характерных особенностей машин этой категории.

Вот. Ну, в общем, так примерно. То есть мы получили нагрузки смысловые, то есть смысл тех или иных тех или иных значений свойств автомобилей. Как влияет цвет автомобиля на стоимость.

Ну, в общем, мы рассмотрели с вами на таком реальном примере уже. Ну, не до конца, может быть, тут ещё кластеры можно было посмотреть. Ну, в общем, так в основном посмотрели, как выглядит решение реальной задачи на реальных данных, имеющие уже, задачи уже имеющие не только учебное значение, но и практическое, и научное значение. Вот. И на этом, значит, наше занятие заканчивается. Но можно вопросы задать, если есть желание, какие-то вопросы, можете задать.

С каждой парой становится всё понятнее.

Ну это правильно, так и должно быть, ребята. А когда вы будете заканчивать, это вообще всё будет понятно. А когда вы получите диплом, вы будете удивляться, как это что-то я там не понимал тогда. Ну это странно так. Вот когда закончишь, кажется, что всё это не так-то и сложно, в общем-то. Ну всё правильно. Значит, вы развиваетесь, значит, вы что-то усваиваете. Я же не зря эти работы вам показываю разноплановые, но как бы идея вот их, подхода к их решению, он сходный. Я хочу вам показать, что в разных предметных областях, для разных целей могут решаться задачи сходными методами. Это то, что обычно вот в науке является очень ценным. Вот математика, она разрабатывает методы, которые применяются в самых разных областях, одни и те же методы. Вот. Если они реализованы в универсальной форме, тогда это вот и становится возможным.

Да, вот кто-то тут исходные данные должны быть на русском. Не обязательно, ребята, на любом языке, но на таком, в котором есть э-э, ну, в котором есть алфавит, который виден в той кодировке нормальной, которая используется в системе ЭИДОС. В IBM 866. Вот. Ну, скажем, вот китайский не проходят, там иероглифы, они не отображаются. Арабский тоже. А вот английский, немецкий, французский, испанский, итальянский там, пожалуйста.

Какие ещё вопросы, ребята?

Вопросов больше нет.

Всё. Следующий вопрос, когда закончится занятие в конце концов, да? Ну всего хорошего, ребята. До свидания. Счастливого.

Спасибо большое, до свидания. Доброго вечера.

Да, до свидания, счастливо.

До свидания.

До свидания.