

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И. Т. ТРУБИЛИНА»**

Факультет прикладной информатики

Кафедра компьютерных технологий и систем

КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии

на тему: АСК-анализ набора данных WEBTOON ORIGINALS на основе
данных портала Kaggle

Выполнил студент группы: ИТ2041 Сулейманов Нургелди Сердарович

Допущен к защите: _____

Руководитель проекта: д.э.н., к. т. н., профессор Луценко Е. В. ()

(подпись, расшифровка подписи)

Защищен 19.02.2022

(дата)

Оценка отлично

Краснодар 2022

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ Федеральное
государственное бюджетное образовательное учреждение высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И. Т. ТРУБИЛИНА»

Факультет прикладной информатики

РЕЦЕНЗИЯ

на курсовую работу

Студента Сулейманов Нургелди Сердарович

курса 2 очной формы обучения группы ИТ2041

Направление подготовки 09.04.02 «Информационные системы и технологии»

Наименование темы «АСК-анализ набора данных WEBTOON ORIGINALS на
основе данных портала Kaggle»

Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор

(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсовой работы

№ п/п	Показатель	Оценка соответствия («неудовлетворительно», «удовлетворительно», «хорошо», «отлично»)
1	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	5
2	Уровень и корректность использования в работе различных методов исследований	5
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	5
4	Ясность, четкость, последовательность и обоснованность изложения	5
5	Применение современных технологий обработки информации	5
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	5
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	5
8	Ответы на вопросы при защите	5

Достоинства работы _____

Недостатки работы: не полностью соответствует шаблону описания курсовой работы 2022 года: <http://ic.kubagro.ru/aidos/The structure of the description of scientific research using ASC-analysis and the Eidos system.doc>

Итоговая оценка при защите
отлично

Рецензент _____ (Е. В. Луценко)

«19» февраля 2022 г.

РЕФЕРАТ

Курсовая работа содержит: 29 страницу, 23 рисунка, 8 литературных источников.

Ключевые слова: СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, AIDOS-X, КОГНИТИВНЫЙ АНАЛИЗ, МОДЕЛИ, ШКАЛЫ, КЛАССЫ.

Целью работы является провести автоматизированный системно-когнитивный анализ набора данных WEBTOON ORIGINALS на основе данных портала Kaggle. Добиться этого можно анализом методов формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами, принятия решений и исследования моделируемой предметной области путем исследования модели.

СОДЕРЖАНИЕ

Содержание.....	4
ВВЕДЕНИЕ.....	5
1 СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ.....	6
1.1 Описание решения.....	6
1.2 Преобразование исходных данных из CSV-формата в файл формата XLSX	8
1.3 Ввод выборки в систему AIDOS-X	10
1.4 Синтез и верификация статистических и интеллектуальных моделей ..	13
1.5 Результаты верификации моделей	15
2 РЕШЕНИЕ ЗАДАЧ НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	19
2.1 Решение задачи идентификации.....	19
2.2 Кластерно-конструктивный анализ	22
2.3 Нелокальные нейронные сети и нейтроны	24
2.4 SWOT и REST-матрицы и диаграммы	27
ЗАКЛЮЧЕНИЕ	30
СПИСОК ЛИТЕРАТУРЫ.....	31

ВВЕДЕНИЕ

Современные информационные системы получают все большее развитие благодаря технологиям искусственного интеллекта. Оценка качества математических моделей некоторых из них не выносит критики. В данной курсовой работе рассмотрено решение задачи АСК-анализа WEBTOON ORIGINALS на основе данных портала Kaggle.

Целью данной курсовой работы является.

Задачами, поставленными в данной курсовой работе, являются:

- подготовка исходных данных и формализация предметной области;
- синтез и верификация статистических и системно-когнитивных моделей и выбор наиболее достоверной модели;
- решение различных задач в наиболее достоверной модели: прогнозирование, поддержка принятия решений, исследование полученных моделей.

Объектом исследования данной работы является выборка данных об WEBTOON ORIGINALS.

Результатом данной работы можно считать получение теоретических и практических знаний в области анализа работы систем искусственного интеллекта и анализа результата их работы.

Курсовая работа состоит из введения, двух глав, заключения и списка литературы. Общий объем работы составляет 31 страницу.

1 СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1 Описание решения

В качестве метода исследования, решения проблемы и достижения цели было решено использовать новый метод искусственного интеллекта:

Автоматизированный системно-когнитивный анализ или АСК-анализ. Основной причиной выбора АСК-анализа является то, что он включает теорию и метод количественного выявления в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал и единицах измерения.

Очень важным является также то, что АСК-анализ имеет свой развитый доступный программный инструментарий, в качестве которого в настоящее время выступает Универсальная когнитивная аналитическая система Aidos-X. Система «Эйдос» выгодно отличается от других интеллектуальных систем следующими параметрами:

- разработана в универсальной постановке, не зависящей от предметной области. Поэтому она является универсальной и может быть применена во многих предметных областях, в которых не требуется автоматического, т. е. без непосредственного участия человека в реальном времени решения задач идентификации, прогнозирования, принятия решений и исследования предметной области;
- находится в полном открытом бесплатном доступе причем с актуальными исходными текстами;
- является одной из первых отечественных систем искусственного интеллекта персонального уровня, т. е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта: «имеет нулевой порог входа»;

- обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения (т.е. не предъявляет жестких требований к данным, которые невозможно выполнить, а обрабатывает те данные, которые есть);

- содержит большое количество локальных (поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений;

- поддерживает on-line среду накопления знаний и широко используется во всем мире;

- обеспечивает мультиязычную поддержку интерфейса на 51 языке.

Языковые базы входят в инсталляцию и могут пополняться в автоматическом режиме;

- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA);

- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение с использованием этих знаний задач классификации, поддержки принятия решений и исследования предметной области путем исследования ее системно-когнитивной модели, генерируя при этом очень большое количество табличных и графических выходных форм (развития когнитивная графика), у многих из которых нет никаких аналогов в других системах (примеры форм можно посмотреть в работе;

– хорошо имитирует человеческий стиль мышления: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции;

– вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторности всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ) предлагает без какой-либо предварительной обработки осмыслить эти данные и тем самым преобразовать их в информацию, а затем преобразовать эту информацию в знания путем ее применения для достижения целей (т.е. для управления) и решения задач классификации, поддержки принятия решений и содержательного эмпирического исследования моделируемой предметной области.

1.2 Преобразование исходных данных из CSV-формата в файл формата XLSX

С электронного ресурса [kaggle.com](https://www.kaggle.com) возьмем набор данных «Набор данных WEBTOON ORIGINALS»,

<https://www.kaggle.com/iridazzle/webtoon-originals-datasets>

CSV-файл содержит 11 столбцов с данными:

- Title – Наименование;
- Genre – Жанр;
- Authors – Автор;
- Weekdays – Неделя;
- Length – Длина;
- Subscribers – Подписчики;
- Rating – Рейтинг;
- Views – Просмотры;
- Likes – Лайки;
- Status – Статус;

– Daily_pass – Ежедневный просмотр.

Для загрузки модели в систему AIDOS-X необходимо конвертировать CSV-файл в файл формата XLSX. Для конвертации был использован онлайн конвертор: <https://convertio.co/ru/csv-xlsx/>

После конвертации необходимо добавить еще один столбец, который будет называться классифицирующим, таким столбцом было решено выбрать title, итоговую таблицу можно увидеть на рисунке 1.

	A	B	C	D	E	F	G	H	I	J	K
1	title	genre	authors	weekdays	length	subscriber	rating	views	likes	status	daily_pass
2	Regina Rena: To the L	FANTASY	Seol Dong	TUESDAY	3	264290	9,71	924069	109080	ONGOING	False
3	Love and Leashes	ROMANCE	Winter	FRIDAY	6	120767	8,3	1096191	121208	ONGOING	True
4	My In-Laws are Obses	FANTASY	seungu,H	SUNDAY	4	184145	9,79	755948	94671	ONGOING	False
5	Manager Kim	ACTION	Jeongjong	SATURDA	3	85150	9,8	260075	32239	ONGOING	False
6	Eternals: The 500 Year	SUPER_HERO	Marvel	FRIDAY	7	175719	7,29	828927	62793	ONGOING	False
7	Tomorrow	DRAMA	Llama	THURSDA	4	57197	9,59	151074	18469	ONGOING	False
8	Hwarang: Flower Knig	ACTION	hwieum,s	FRIDAY	4	68891	9,31	176532	21109	ONGOING	False
9	A DeadbeAT's Meal	SLICE_OF_LIFE	CHEESE	WEDNESD	4	98788	9,24	488273	61064	ONGOING	False
10	A Chance At Last	ROMANCE	Dindon,Se	WEDNESD	4	125242	9,82	465247	59094	ONGOING	False
11	Daytime Star	DRAMA	Chaeun,G	TUESDAY	4	105674	9,72	397045	51320	ONGOING	False
12	Psychic Studies	ACTION	PARK EUN	MONDAY	4	87237	8,54	360454	40501	ONGOING	False
13	The Retreats	ACTION	Hyungwo	SUNDAY	5	35019	8,95	126201	12069	ONGOING	False
14	Discovering Love: A H	ROMANCE	Various Ai	MONDAY,	10	59536	8,99	293781	30216	ONGOING	False
15	Secretary Out-of-Ord	ROMANCE	QuornQu	TUESDAY	8	228969	8,83	1835205	204670	ONGOING	False
16	Our House	SLICE_OF_LIFE	Jinha,Jin J	WEDNESD	9	75575	9,3	465630	65514	ONGOING	False
17	WEBTOON x BTS	FANTASY	WEBTOON	MONDAY,	7	798567	9,73	9427023	1197835	ONGOING	False
18	To Not Die	DRAMA	Parae,Jing	TUESDAY	8	116502	9,58	816533	84610	ONGOING	False
19	THE STAR SEEKERS	FANTASY	HYBE	MONDAY	5	796233	9,84	2303429	293300	ONGOING	False
20	DARK MOON: THE BLO	FANTASY	HYBE	SUNDAY	5	1055574	9,85	5661458	499996	ONGOING	False
21	7FATES: CHAKHO	FANTASY	HYBE	SATURDA	6	1430708	9,86	18572875	1079005	ONGOING	False
22	Murim RPG Simulatio	ACTION	Hyung Gei	FRIDAY	8	116551	9,5	830387	76301	ONGOING	False
23	QUESTISM	FANTASY	TAEWAN,	WEDNESD	8	94106	9,49	677040	78504	ONGOING	False
24	100	ACTION	Hyeong Et	SUNDAY	5	102082	9,17	429338	40749	ONGOING	False
25	Night of Silence	THRILLER	Dongwoo	SUNDAY	5	64599	9,12	230255	21491	ONGOING	False
26	The RUNWAY	ROMANCE	Domi,Nok	THURSDA	12	377452	9,42	4695580	419388	ONGOING	False
27	The Lone Necromanc	FANTASY	JJJ,KIM KY	SUNDAY	10	363989	9,69	4891190	462058	ONGOING	False
28	Bitten Contract	ROMANCE	sungeun	FRIDAY	12	342623	9,65	3707983	388559	ONGOING	False
29	Return of the Mad De	HISTORICAL	JP,Ihy,Yu J	WEDNESD	12	190781	9,76	1676894	202361	ONGOING	False
30	Cierra	ROMANCE	Rana,Moo	TUESDAY	12	295019	9,62	2948677	373607	ONGOING	False
31	Double Click	SPORTS	KIM JANG	SATURDA	11	172643	9,74	1685186	189403	ONGOING	False
32	Our Beloved Summer	ROMANCE	Han Kyoui	FRIDAY	12	308656	9,85	2629739	318470	ONGOING	False
33	Sarah, I'm Sorry: Welc	TIPTOON	MiTl,GUGI	WEDNESD	5	124973	8,41	513720	63742	ONGOING	False
34	Rewriting Extinction	COMEDY	Various Ai	FRIDAY	29	94944	9,72	1405611	254876	ONGOING	False
35	Our Time	DRAMA	seaweed	FRIDAY	17	120671	9,72	1256480	195689	ONGOING	False
36	Hellbound	THRILLER	Choi Gyu	TUESDAY	17	195638	9,33	1652583	127357	ONGOING	False
37	Reincarnated as an Ur	ACTION	Jeon mad	WEDNESD	16	162320	9,65	1770614	211087	ONGOING	False
38	The Greatest Estate D	FANTASY	Lee hyunr	SUNDAY	15	205372	9,88	2701364	377136	ONGOING	False
39	Study Group	ACTION	Hyunmwul	SUNDAY	16	318416	9,76	3422347	416657	ONGOING	False

Рисунок 1 – Фрагмент обучающей выборки

1.3 Ввод выборки в систему AIDOS-X

Для импорта обучающей выборки в систему AIDOS-X необходимо скопировать ее в папку Inp_data и переименовать в Inp_data.xlsx, после этого можно запустить саму программу и универсальный программный интерфейс импорта данных в систему (режим 2.3.2.2), результат заполнения которого представлен на рисунке 2.

2.3.2.2. Универсальный программный интерфейс импорта данных в систему "ЭЙДОС-X++"

Автоматическая формализация предметной области: генерация классификационных и описательных шкал и градаций, а также обучающей и распознаваемой выборки на основе базы исходных данных: "Inp_data"

Задайте тип файла исходных данных: "Inp_data":

- XLS - MS Excel-2003
- XLSX- MS Excel-2007(2010)
- DBF - DBASE IV (DBF/NTX)
- CSV - CSV => DBF конвертер

Стандарт XLS-файла

Стандарт DBF-файла

Стандарт CSV-файла

Задайте параметры:

- Нули и пробелы считать ОТСУТСТВИЕМ данных
- Нули и пробелы считать ЗНАЧЕНИЯМИ данных
- Создавать БД средних по классам "Inp_davr.dbf"?

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:

Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:

Конечный столбец описательных шкал:

Задайте режим:

- Формализации предметной области (на основе "Inp_data")
- Генерации распознаваемой выборки (на основе "Inp_rasp")

Задайте способ выбора размера интервалов:

- Равные интервалы с разным числом наблюдений
- Разные интервалы с равным числом наблюдений

Задание параметров формирования сценариев или способа интерпретации текстовых полей "Inp_data":

- Не применять сценарный метод АСК-анализа
- Применить сценарный метод АСК-анализа
- Применить спец. интерпретацию текстовых полей классов
- Применить спец. интерпретацию текстовых полей признаков

Параметры интерпретации значений текстовых полей "Inp_data":

В качестве классов рассматриваются:

- Значения полей целиком
- Элементы значений полей - слова > символов:
- Элементы значений полей - символы

Выделять уникальные значения и сортировать

Не выделять уникальных значений и не сортировать

В качестве признаков рассматриваются:

- Значения полей целиком
- Элементы значений полей - слова > символов:
- Элементы значений полей - символы

Проводить лемматизацию

Не проводить лемматизацию

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

- Только интервальные числовые значения (например: "1/3-{59873.0000000, 178545.6666667}")
- Только наименования интервальных числовых значений (например: "Минимальное")
- И интервальные числовые значения, и их наименования (например: "Минимальное: 1/3-{59873.0000000, 178545.6666667}")

Ok Cancel

Рисунок 2 – Интерфейс импорта

Следует выделить следующие настройки:

- Тип файла – XLSX;
- Классификационная шкала – 2;
- Описательные шкалы – 3-8.

После этого приложение просит задать размерности модели системы, изменяем до 5 (рисунок 3).

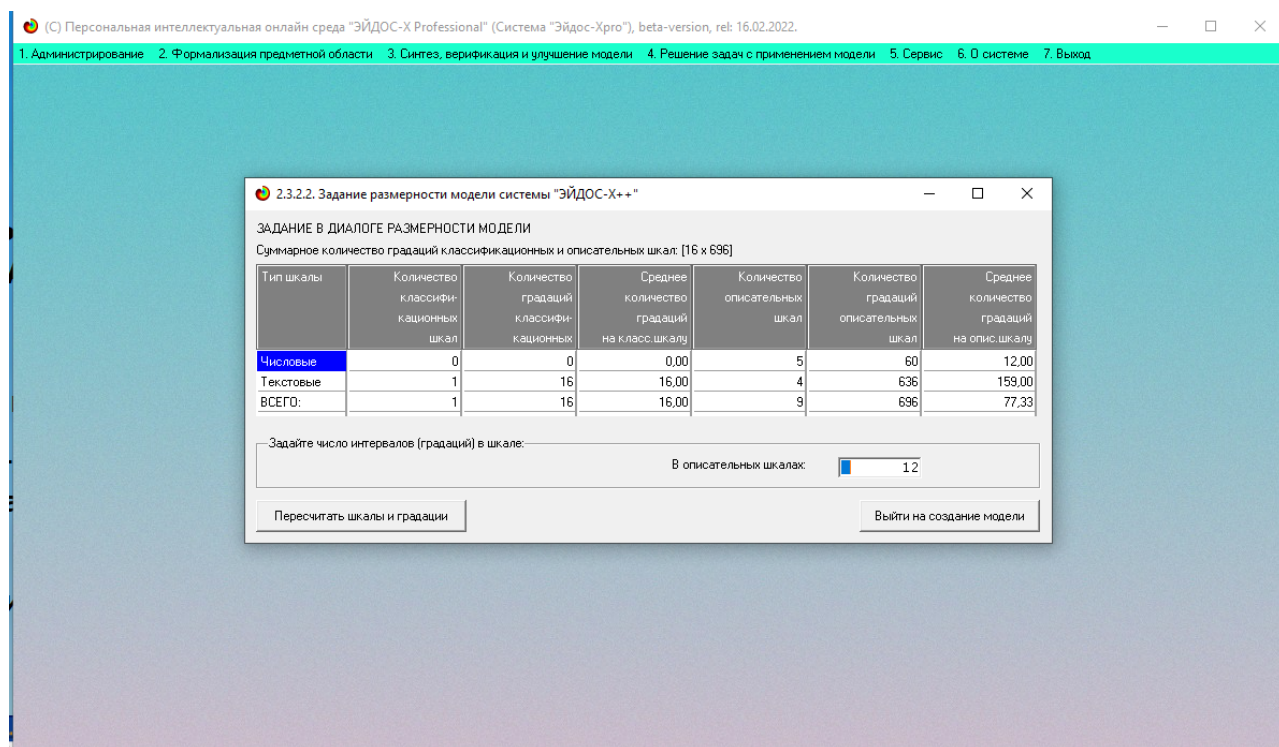


Рисунок 3 – Задание размерностей системы

Процесс импорта данных из внешнего файла в систему представлен на рисунке 4.

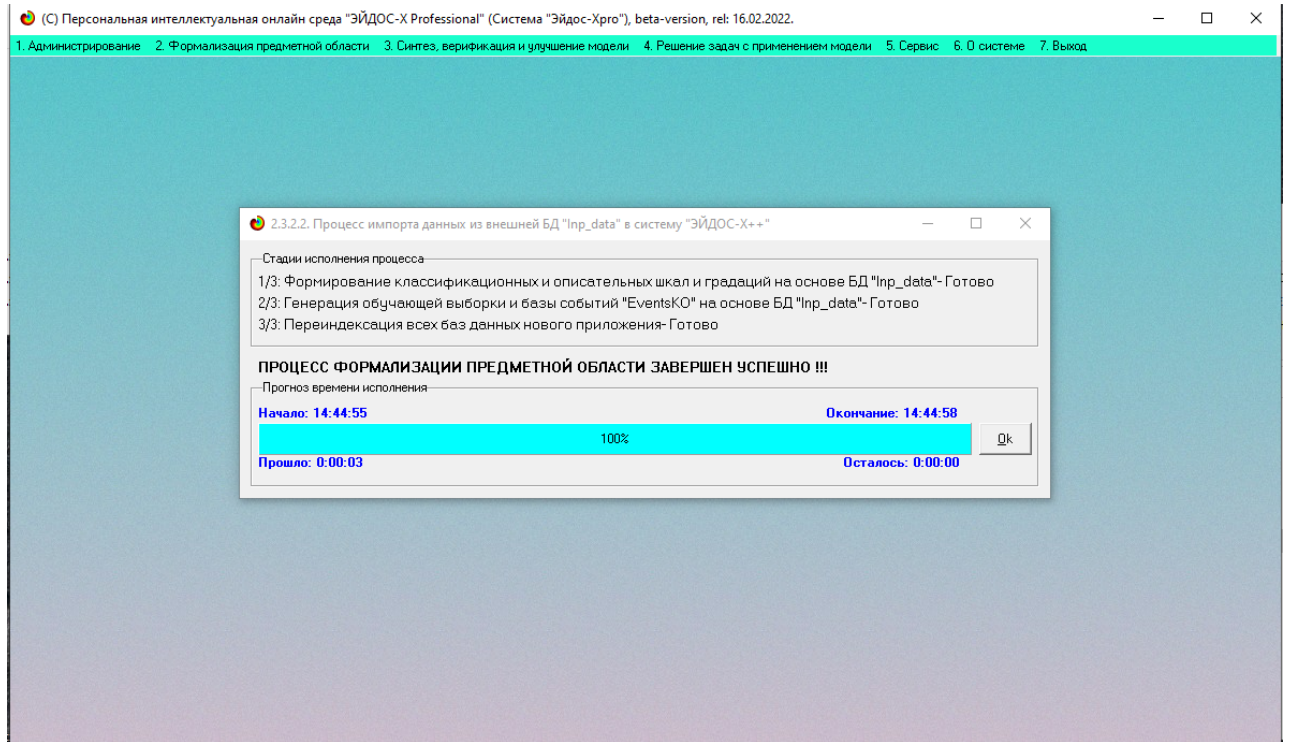


Рисунок 4 – Импорт данных

После загрузки данных система автоматически нашла классификационные шкалы, которые можно посмотреть в режиме 2.1 (рисунок 5) и описательные шкалы, которые можно посмотреть в режиме 2.2 (рисунок 6).

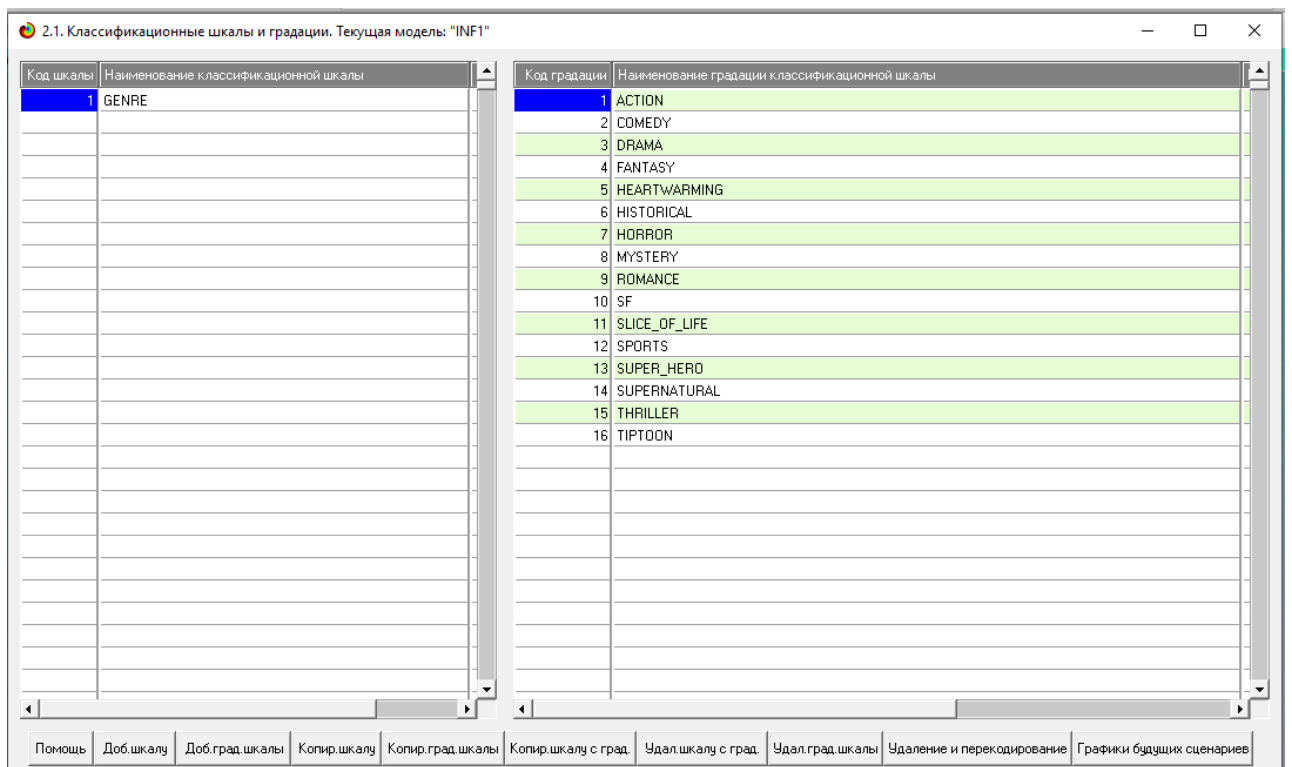


Рисунок 5 – Классификационные шкалы

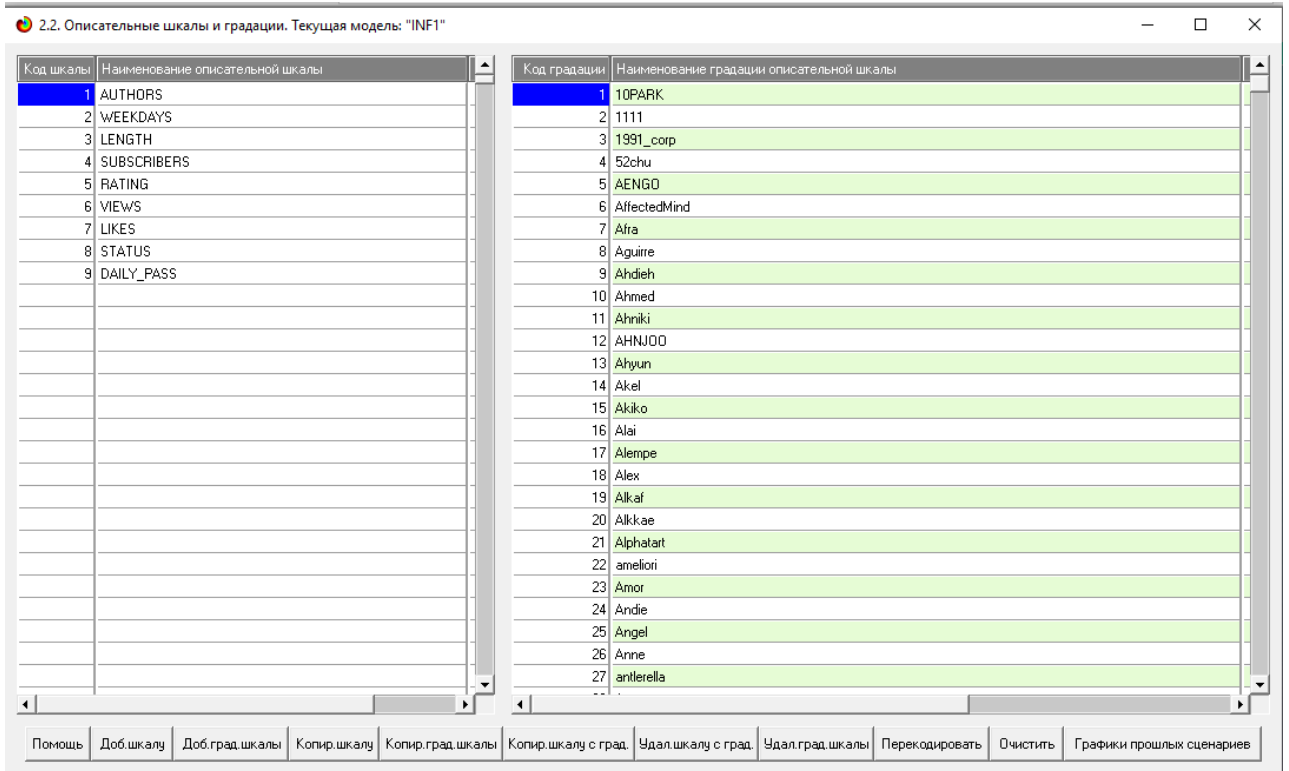


Рисунок 6 – Описательные шкалы

Так же существует возможность ручной корректировки выгруженных данных и добавление новых объектов, которая открывается с помощью режима 2.3.1 (рисунок 7).

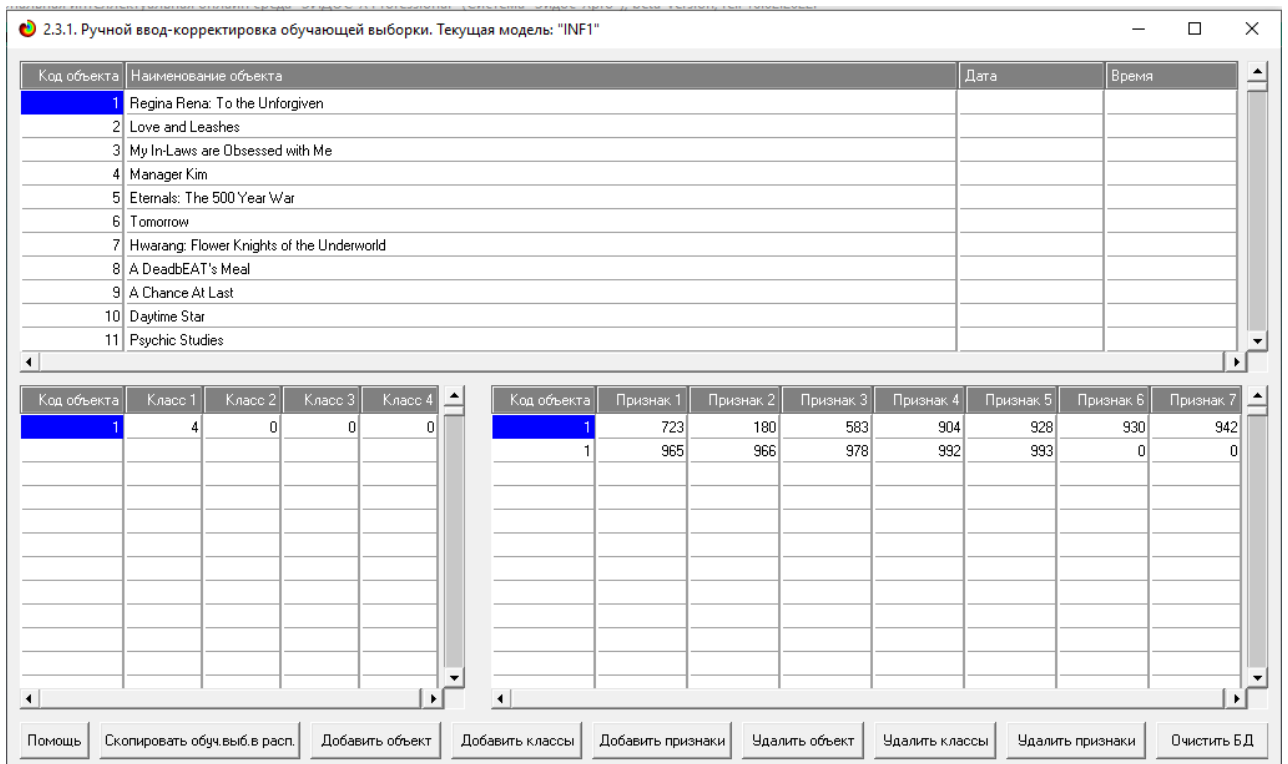


Рисунок 7 – Ручная корректировка

1.4 Синтез и верификация статистических и интеллектуальных моделей

Для синтеза и верификации моделей создан режим 3.5, после его запуска задается модель, которая помечается текущей (рисунок 8).

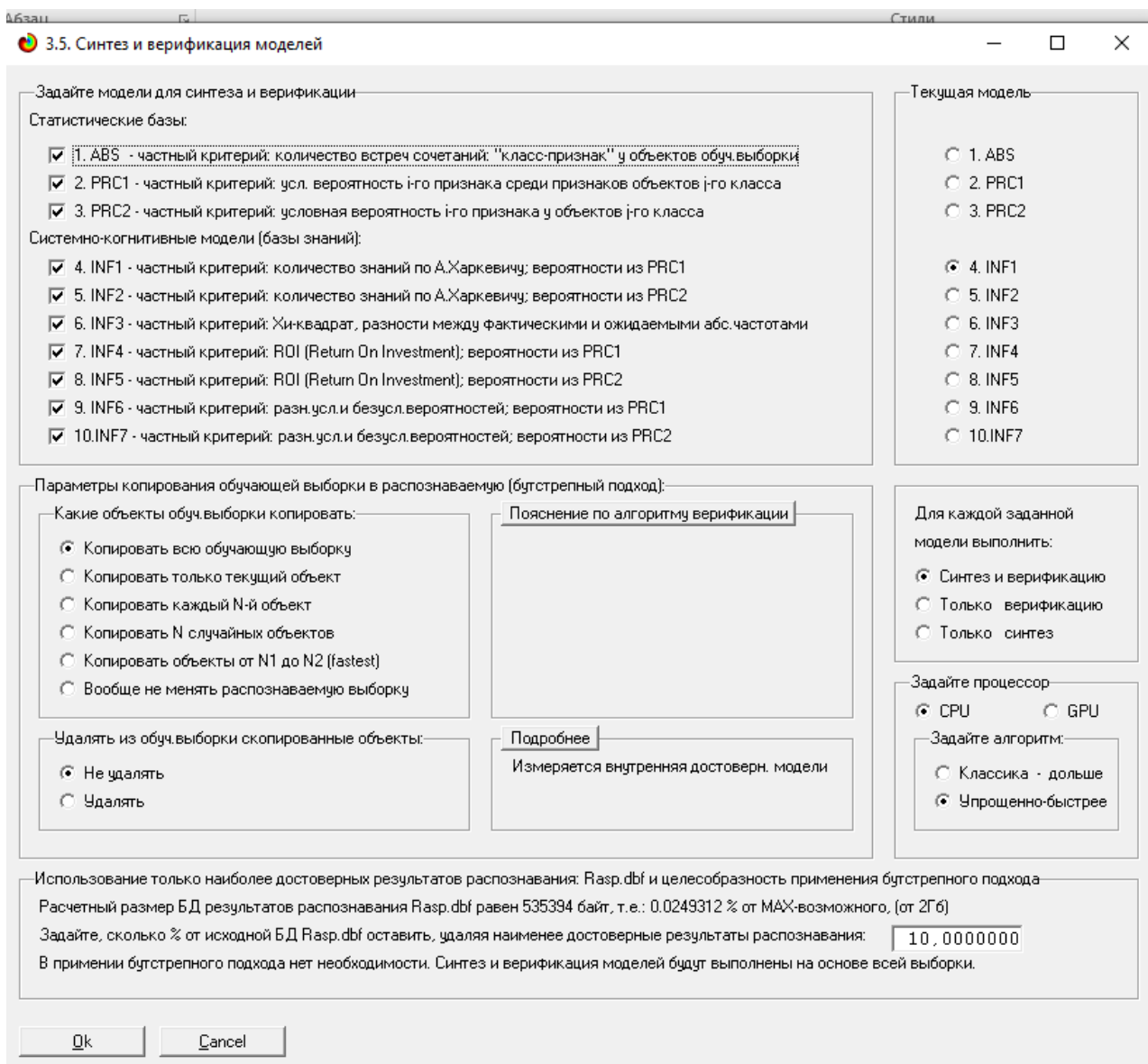


Рисунок 8 – Режим синтеза моделей

Данный режим содержит множество различных методов верификации, но мы используем параметры, которые система предлагает по умолчанию. Стадия процесса исполнения режима 3.5 и прогноз времени исполнения показаны на рисунке 9.

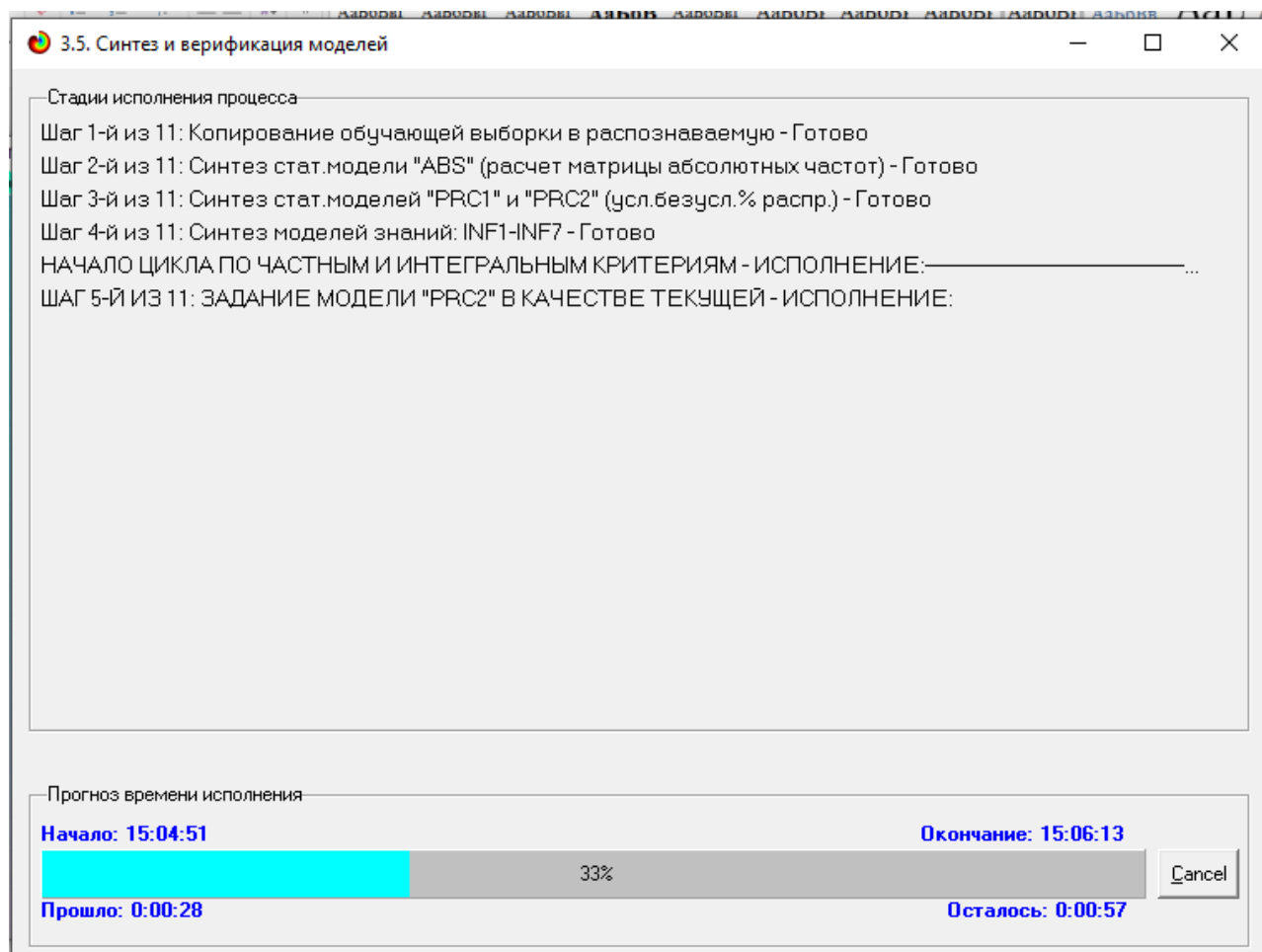


Рисунок 9 – Исполнение режима синтеза моделей

Следует заметить, что синтез и верификация всех моделей занял 16 секунд. После данного этапа можно приступить к выбору наиболее достоверной модели.

1.5 Результаты верификации моделей

С результатами верификации моделей, отличающихся частными критериями, можно ознакомиться в режиме 3.4 системы «Эйдос», они представлены на рисунке 10.

3.4. Обобщ. форма по достов. моделей при разн. крит. Текущая модель: "INF5"

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма уровней сход. истинно-поло. решений (STR)	Сумма уровней сход. истинно-отриц. решений (STN)	Сумма уровней ложно-поло. решений
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс частот с обр...	669	669		1003		0.400	1.000	0.572	474.326		751
1.ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс частот по признак...	669	669		1003		0.400	1.000	0.572	301.066		282
2. PRC1 - частный критерий: усл. вероятность н-го признака сред...	Корреляция усл отн частот с о...	669	669		1003		0.400	1.000	0.572	474.326		751
2.PRC1 - частный критерий: усл. вероятность н-го признака сред...	Сумма усл отн частот по приз...	669	669		1003		0.400	1.000	0.572	323.197		535
3. PRC2 - частный критерий: условная вероятность н-го признака...	Корреляция усл отн частот с о...	669	669		1003		0.400	1.000	0.572	474.326		751
3.PRC2 - частный критерий: условная вероятность н-го признака...	Сумма усл отн частот по приз...	669	669		1003		0.400	1.000	0.572	264.645		444
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	669	497	813	190	172	0.723	0.743	0.733	52.074	52.974	11
4.INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	669	648	563	440	21	0.596	0.969	0.738	108.039	15.651	17
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	669	499	798	205	170	0.709	0.746	0.727	48.391	47.337	15
5.INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	669	647	509	494	22	0.567	0.967	0.715	112.224	15.954	21
6. INF3 - частный критерий: Хинкварадг, разности между фактик...	Семантический резонанс зна...	669	468	481	522	201	0.473	0.700	0.564	193.307	126.386	161
6.INF3 - частный критерий: Хинкварадг, разности между фактик...	Сумма знаний	669	468	481	522	201	0.473	0.700	0.564	180.922	106.552	87
7. INF4 - частный критерий: ROI (Return On Investment); веротно...	Семантический резонанс зна...	669	470	993	10	199	0.979	0.703	0.818	53.423	53.400	0
7.INF4 - частный критерий: ROI (Return On Investment); веротно...	Сумма знаний	669	655	447	556	14	0.541	0.979	0.697	35.972	0.572	2
8. INF5 - частный критерий: ROI (Return On Investment); веротно...	Семантический резонанс зна...	669	470	987	16	199	0.967	0.703	0.814	53.591	53.068	0
8.INF5 - частный критерий: ROI (Return On Investment); веротно...	Сумма знаний	669	652	410	593	17	0.524	0.975	0.681	30.988	0.496	2
9. INF6 - частный критерий: разн усл и безуслов вероятностей; вер...	Семантический резонанс зна...	669	459	472	531	210	0.464	0.686	0.553	171.621	119.800	161
9.INF6 - частный критерий: разн усл и безуслов вероятностей; вер...	Сумма знаний	669	468	467	536	201	0.466	0.700	0.559	61.954	40.544	87
10.INF7 - частный критерий: разн усл и безуслов вероятностей; ве...	Семантический резонанс зна...	669	491	435	568	178	0.464	0.734	0.568	147.842	91.313	151
10.INF7 - частный критерий: разн усл и безуслов вероятностей; ве...	Сумма знаний	669	503	406	597	166	0.457	0.752	0.569	40.263	22.090	65

Помощь по мерам достоверности Помощь по частотным распределениям TP/TN,FP/FN (TP-FP)/(TN-FN) (T-F)/(T+F-F100) Задать интеграл сглаживания

Рисунок 10 – Оценки достоверности моделей

Для оценки достоверности моделей в АСК-анализе и систему Aidos-X используется F-мера Ван Ризбергера и L-мера, представляющая собой ее нечеткое мультиклассовое обобщение, предложенное профессором Луценко (рисунок 11). Наиболее достоверной оказалась модель **INF4**. Точность данной модели по F-мере Ван Ризбергера составляет **0.818**, а по L1-мере профессора Луценко – **0.969**, а $L2=0,954$. L1-мера, предложенная профессором Луценко является более достоверной, в сравнении с F-мерой Ван Ризбергера. Таким образом, уровень достоверности прогнозирования с применением модели выше экспертных оценок, достоверность которых считается равной примерно 70%.

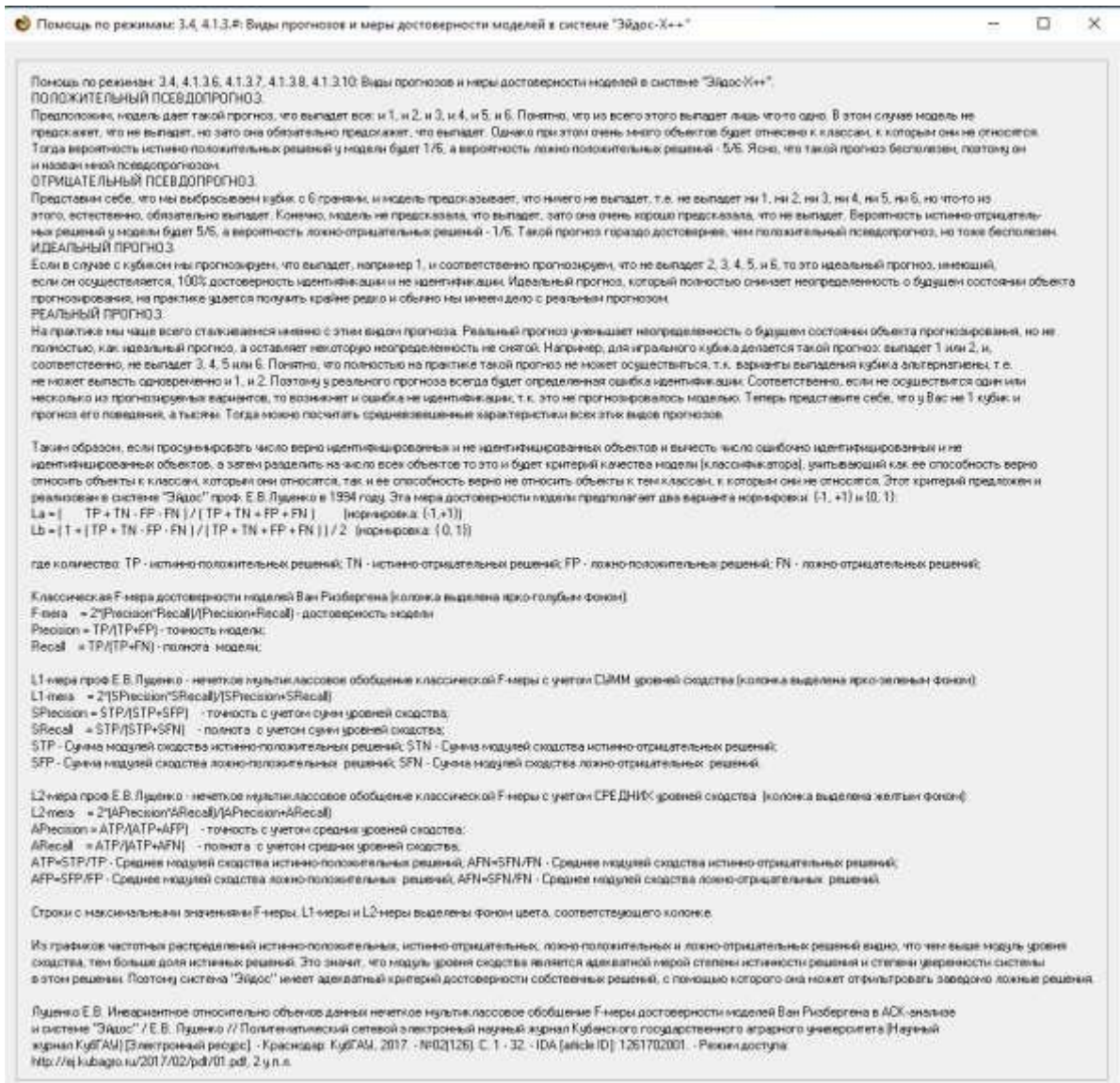


Рисунок 11 – Виды прогнозов и меры достоверностей в системе «Эйдос»

Также необходимо обратить внимание на то, что статистические модели, как правило, дают более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, и практически никогда – более высокую. Этим и оправдано применение моделей знаний из 23 интеллектуальных технологий. На рисунке 12 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и не идентифицированных ситуаций в наиболее достоверной модели INF3.

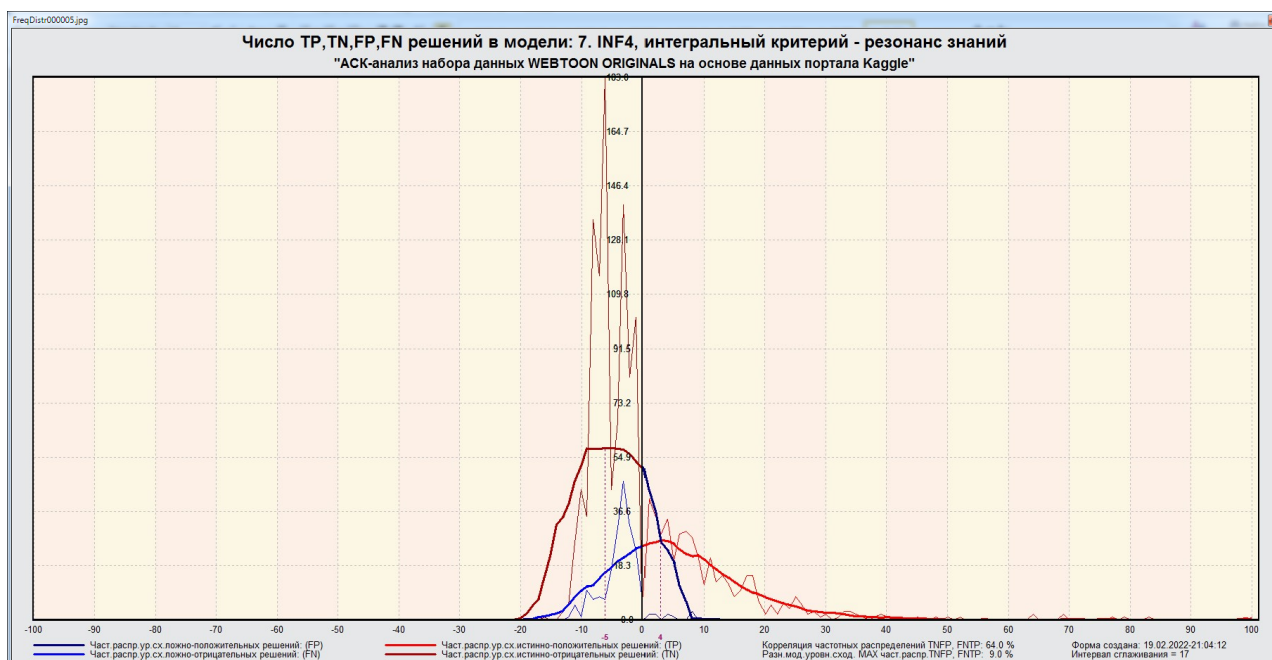


Рисунок 12 – Частотные распределения верно и ошибочно идентифицированных и не идентифицированных состояний объекта моделирования в зависимости от уровня сходства в модели INF3

Левое распределение, включает только истинно-отрицательные значения, а правое включает ложные отрицательные и истинно-положительные решения. Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу идентификации видов героев по его характеристикам и другие задачи.

Для положительных решений от 0% до примерно 10% количество ложных решений больше числа истинных, но далее идет на спад и ложные решения перестают встречаться после 10%.

2 РЕШЕНИЕ ЗАДАЧ НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1 Решение задачи идентификации

В соответствии с технологией АСК-анализа необходимо задать текущую модель в режиме 5.6, в ее качестве выберем наиболее достоверную модель INF3, что показано на рисунке 13.

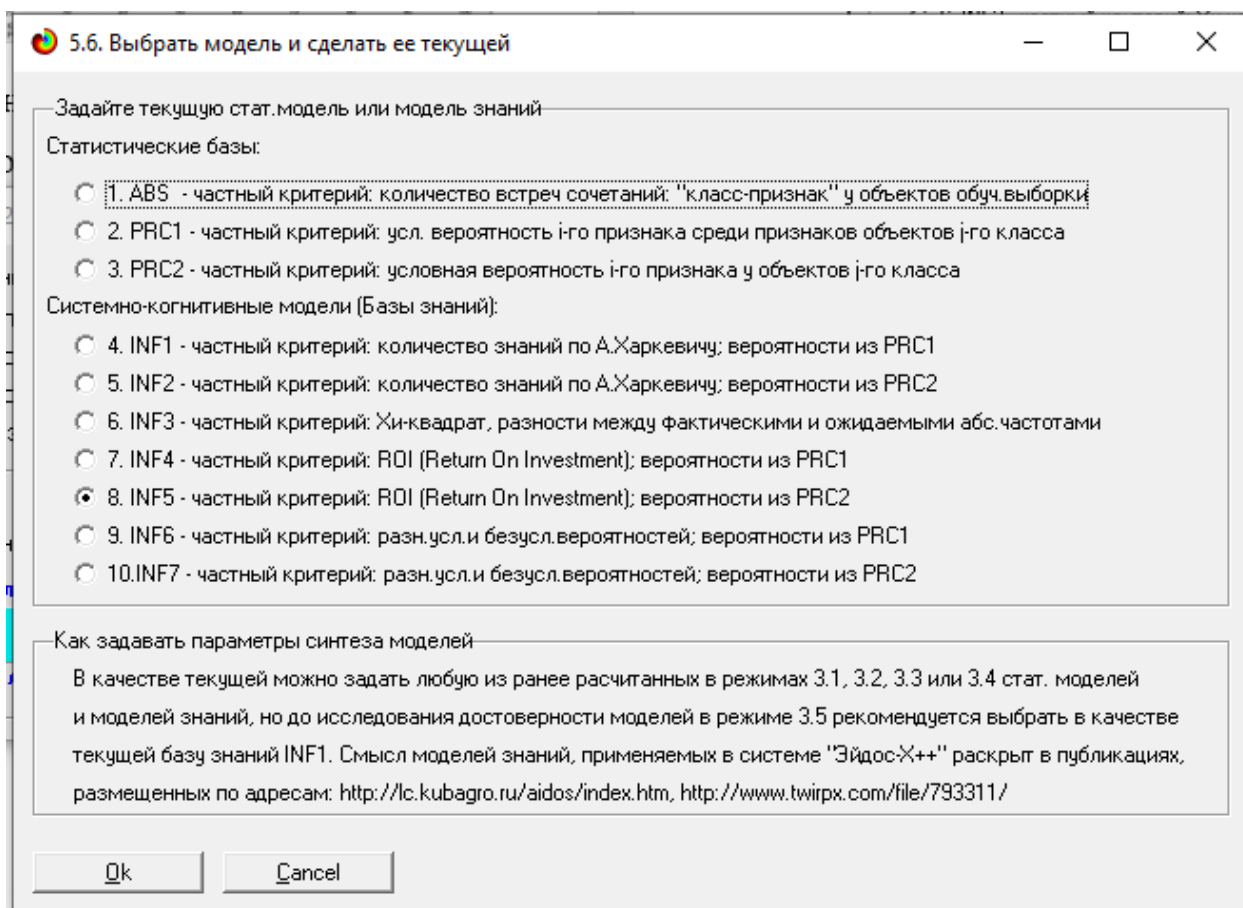


Рисунок 13 – Выбор текущей модели

После этого необходимо провести пакетное распознавание текущей модели в режиме 4.2.1 (рисунок 14)

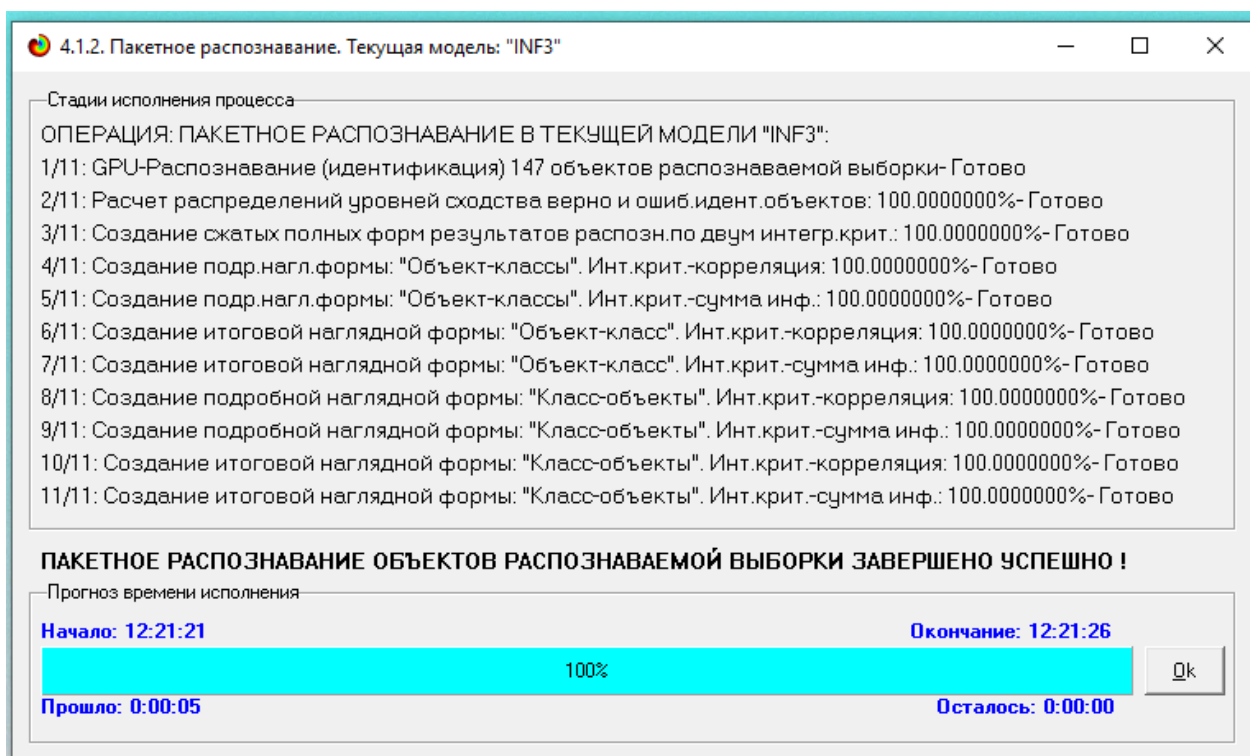


Рисунок 14 – Пакетное распознавание текущей модели

Режим 4.1.3 системы «Эйдос» обеспечивает отображение результатов идентификации и прогнозирования в различных формах:

- Подробно наглядно: "Объект – классы";
- Подробно наглядно: "Класс – объекты";
- Итоги наглядно: "Объект – классы";
- Итоги наглядно: "Класс – объекты";
- Подробно сжато: "Объект – классы";
- Обобщенная форма по достоверности моделей при разных интегральных критериях;
- Обобщенный статистический анализ результатов идентификации по моделям и интегральным критериям;
- Статистический анализ результатов идентификации по классам, моделям и интегральным критериям;
- Распознавание уровня сходства при разных моделях и интегральных критериях;

– Достоверность идентификации классов при разных моделях и интегральных критериях.

Рассмотрим подробнее режим работы 4.2.2.3, результаты которого представлены на рисунке 15.

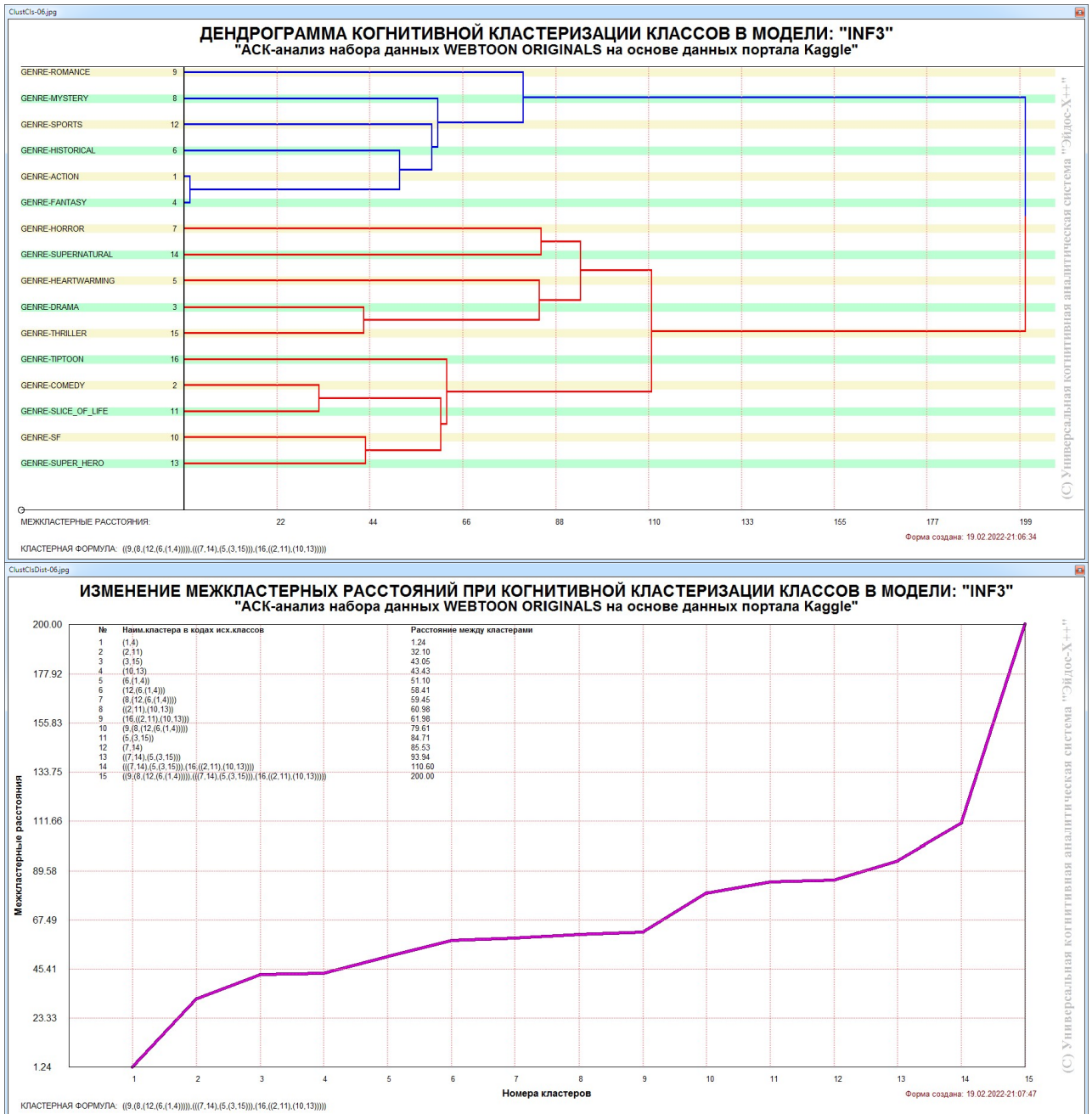
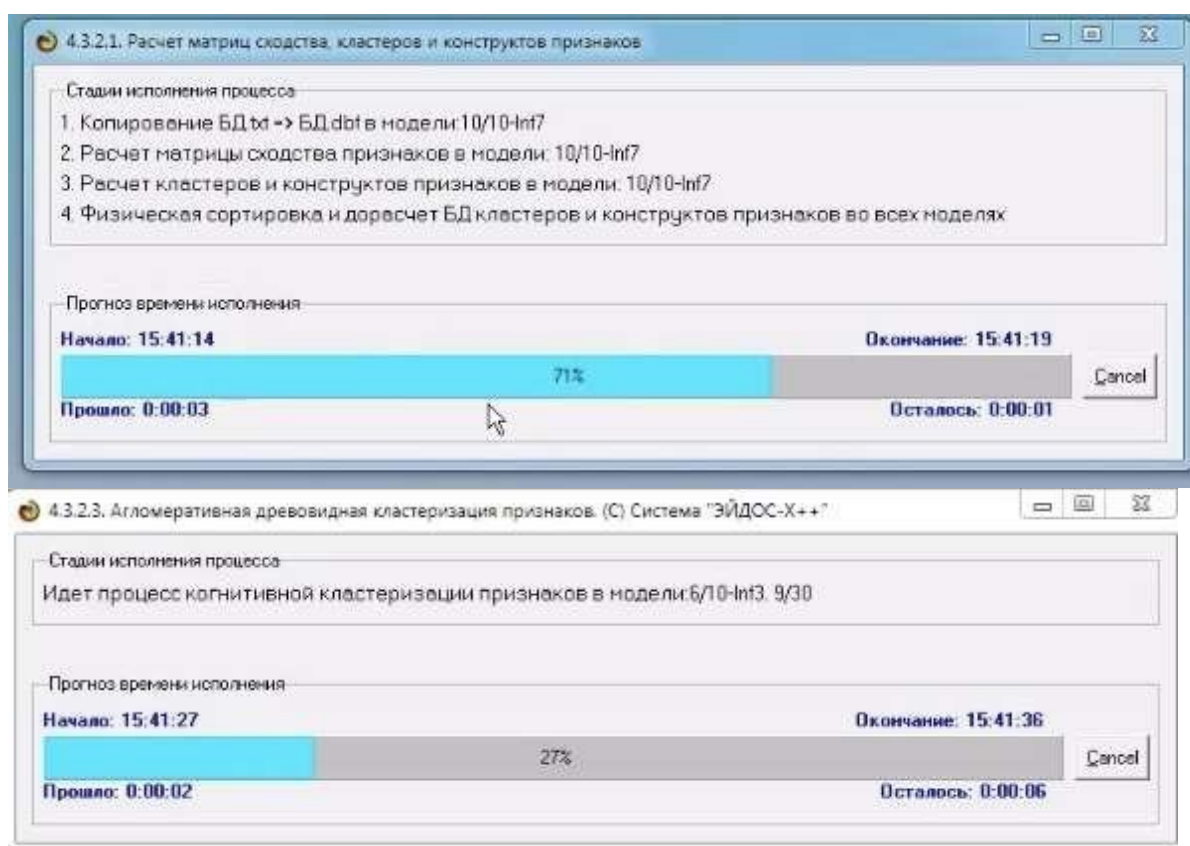


Рисунок 15 – Результат режима работы 4.2.2.3, дендрограмма и график изменения межкластерных расстояний при когнитивной кластеризации классов в модели INF3.

Из рисунков выше видно, что результаты идентификации являются динамичными.

2.2 Кластерно-конструктивный анализ

Для выявления сходства-различия обобщенных образов различных результатов научной деятельности по характерных для них системам значений показателей можно осуществить с помощью режимов 4.3.2.1 и 4.3.2.3, результаты выполнения этих режимов показаны на рисунке 17 и 18.



2.3 Нелокальные нейронные сети и нейтроны

После Каждому классу системно-когнитивной модели соответствует нелокальный нейрон, совокупность которых образует нелокальную нейронную сеть.

На рисунке 19 изображено графическое отображение нелокальных нейронов в системе Aidos-X.

4.4.10.Графическое отображение нелокального нейрона в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	GENRE-ACTION
2	GENRE-COMEDY
3	GENRE-DRAMA
4	GENRE-FANTASY
5	GENRE-HEARTWARMING
6	GENRE-HISTORICAL
7	GENRE-HORROR
8	GENRE-MYSTERY
9	GENRE-ROMANCE
10	GENRE-SCIENCE_FICTION

Подготовка визуализации нейрона: 1 "GENRE-ACTION" в модели: 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
992	STATUS-ONGOING	15.899
965	RATING-12/12-{9.5241667, 9.9500000}	6.140
930	LENGTH-1/12-{1.0000000, 113.0000000}	5.965
993	DAILY_PASS-False	4.741
966	VIEWS-1/12-{56532.0000000, 16949688.4166667}	3.667
361	AUTHORS-Jeho	2.718
322	AUTHORS-Hyeong	1.812
456	AUTHORS-Kwangsu	1.812
835	AUTHORS-UMKY	1.812
990	AUTHORS-ZUMA	1.812

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
990	STATUS-COMPLETED	-16.129
925	WEEKDAYS-SATURDAY	-7.236
924	WEEKDAYS-MONDAY	-6.861
927	WEEKDAYS-THURSDAY	-5.893
994	DAILY_PASS-True	-4.542
928	WEEKDAYS-TUESDAY	-4.048
931	LENGTH-2/12-{113.0000000, 225.0000000}	-3.885
929	WEEKDAYS-WEDNESDAY	-3.456
964	RATING-11/12-{9.0983333, 9.5241667}	-2.554
970	VIEWS-2/12-{6000000.0000000, 10000000.0000000}	-2.000

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

НЕЙРОН Максимальное количество отображаемых рецепторов: 999 Минимальный вес.коэф.отображаемых рецепторов: 0,000

Сортировать рецепторы: по информативности по модулю информативности

Отображать рецепторы: с наименованиями только с кодами

Рисунок 17 – Нелокальные нейроны

Для каждого технологического фактора в соответствии с предложенной моделью определяется величина и направление его влияния на осуществление всех желаемых и не желаемых хозяйственных ситуаций. Для каждой ситуации эта информация отображается в различных текстовых и графических формах, в частности в форме нелокального нейрона (рисунок 20). На данной диаграмме цвет линии означает знак связи (красный – положительная, синий – отрицательная), а толщина – ее модуль

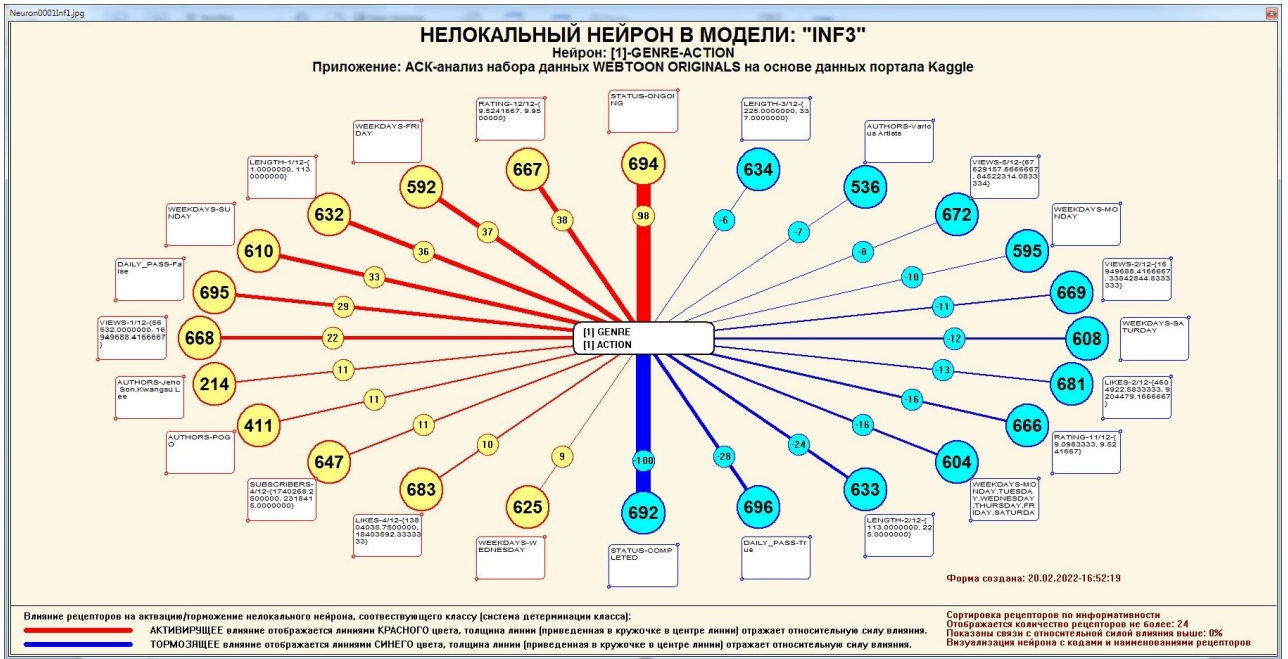


Рисунок 18 – Нелокальный нейрон в модели «INF3»

Дополнение модели нейрона связями факторов позволяет построить классическую когнитивную карту ситуации (будущего состояния АОУ). Детальная внутренняя структура любой связи отображается в форме инвертированной когнитивной диаграммы (рисунок 21). Необходимо отметить, что все указанные графические формы генерируются системой Aidos-X автоматически в соответствии с созданной моделью.

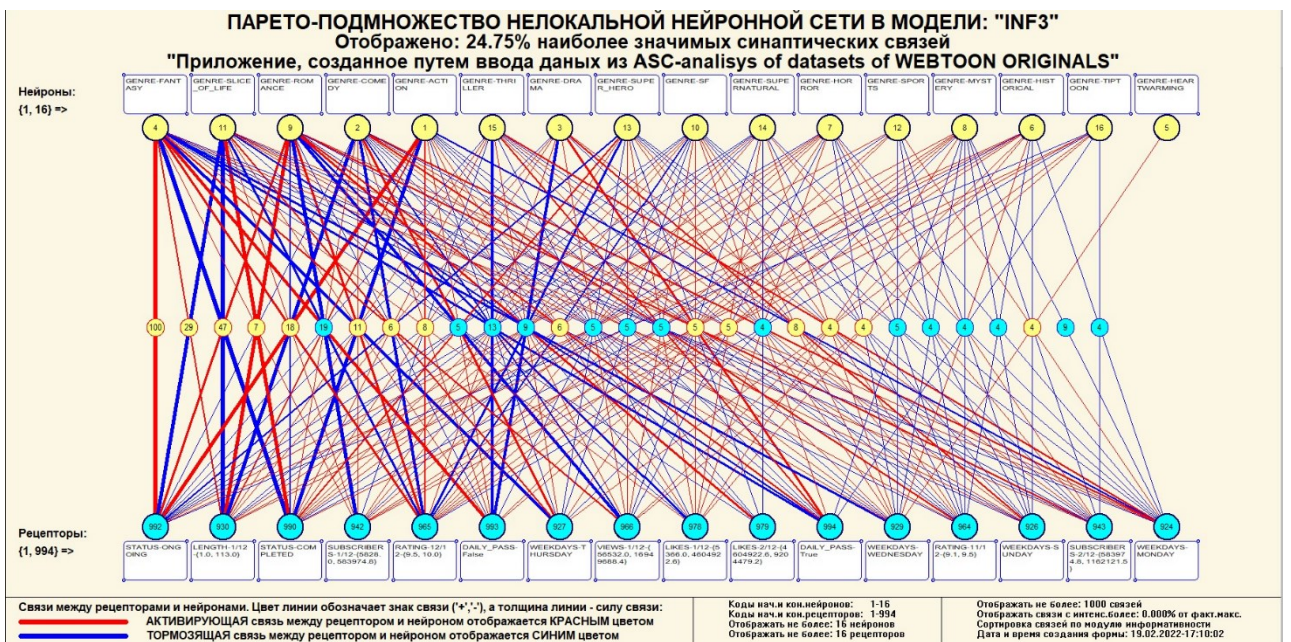


Рисунок 19 – Паретто – подмножество нелокальной нейронной сети

2.4 SWOT и REST-матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают не формализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных. Подобная технология разработана давно, ей уже около 30 лет, но она малоизвестна – это интеллектуальная система Aidos-X. Данная система всегда обеспечивала возможность проведения количественного автоматизированного SWOT-анализа без использования экспертных оценок непосредственно на основе эмпирических данных. Результаты SWOT анализа выводились в форме информационных портретов. В версии системы под MS Windows: Aidos-X++ предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунки 22 и 23).

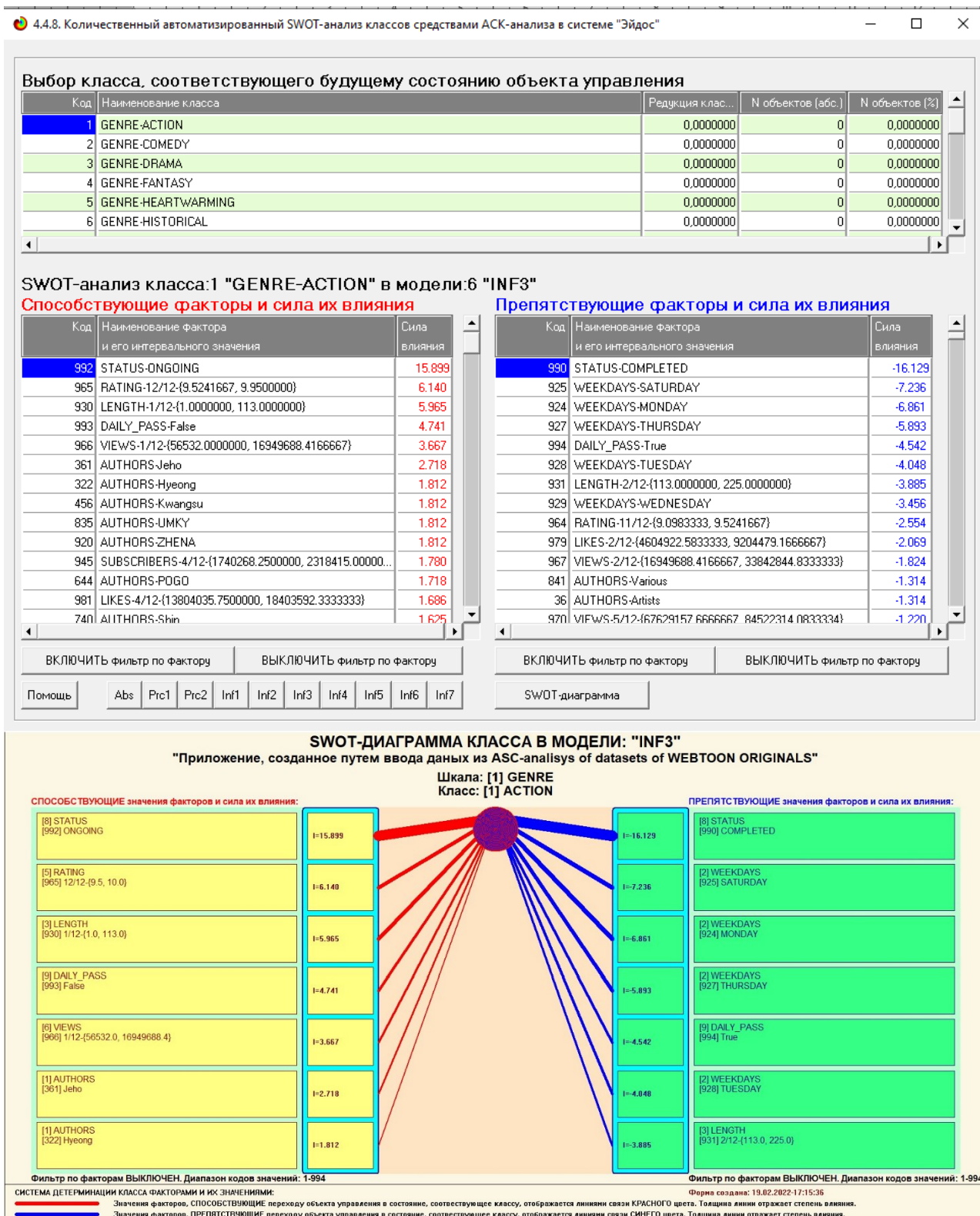


Рисунок 20 – Количественный SWOT-анализ классов в модели PRC2

4.4.9 Количественный автоматизированный SWOT-анализ значений факторов средствами ACK-анализа в системе "Эйдос"

Выбор значения фактора, оказывающего влияние на переход объекта управления в будущие состояния

Код	Наименование значения фактора
1	AUTHORS-10PARK
2	AUTHORS-1111
3	AUTHORS-1991_corp
4	AUTHORS-52chu
5	AUTHORS-AENGO
6	AUTHORS-AffectedMind

SWOT-анализ значения фактора: 1 "AUTHORS-10PARK" в модели: 6 "INF3"

СПОСОБСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влиян...
15	GENRE-THRILLER	0.929

ПРЕПЯТСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
4	GENRE-FANTASY	-0.169
9	GENRE-ROMANCE	-0.153
3	GENRE-DRAMA	-0.112
1	GENRE-ACTION	-0.094
2	GENRE-COMEDY	-0.087
11	GENRE-SLICE_OF_LIFE	-0.080
14	GENRE-SUPERNATURAL	-0.059
10	GENRE-SF	-0.054
13	GENRE-SUPER_HERO	-0.042
7	GENRE-HORROR	-0.031
12	GENRE-SPORTS	-0.015
8	GENRE-MYSTERY	-0.013
16	GENRE-TIPTOON	-0.010
6	GENRE-HISTORICAL	-0.007

ВКЛЮЧИТЬ фильтр по кл.шкале ВЫКЛЮЧИТЬ фильтр по кл.шкале ВКЛЮЧИТЬ фильтр по кл.шкале ВЫКЛЮЧИТЬ фильтр по кл.шкале

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма

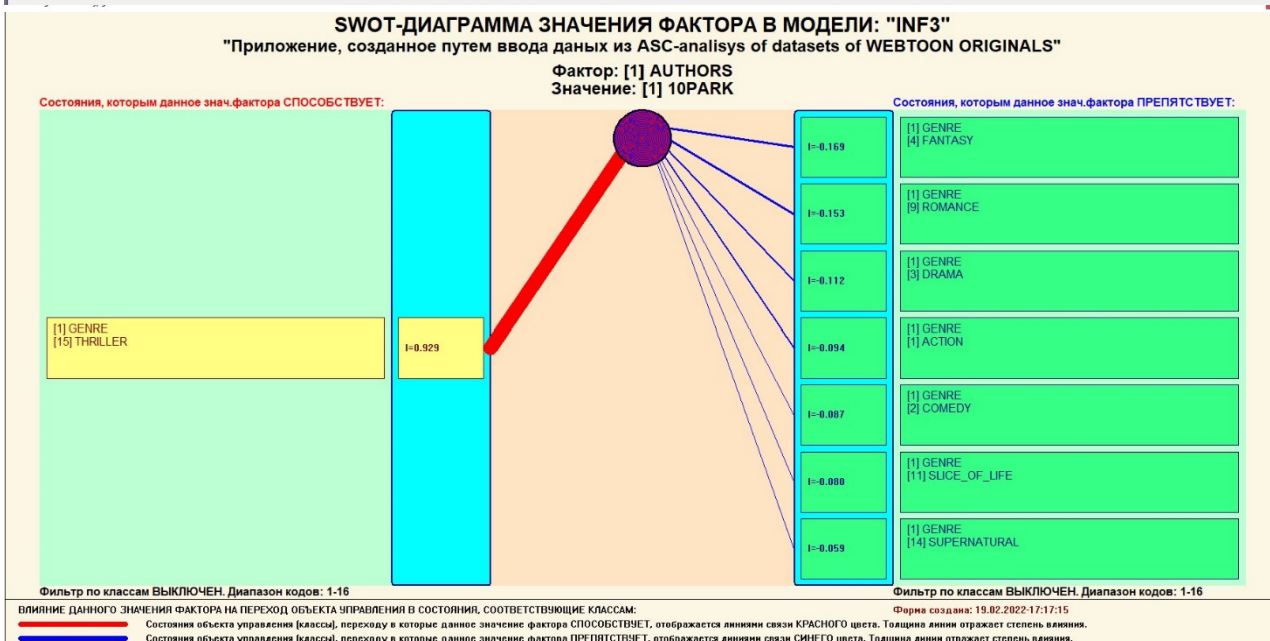


Рисунок 21 – Количественный SWOT-анализ значений в модели PRC2

ЗАКЛЮЧЕНИЕ

Целью работы было провести автоматизированный системно-когнитивный анализ WEBTOON ORIGINALS на основе данных портала Kaggle.

Для этого были изучены методы формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами. Построение моделей было осуществлено с помощью системы искусственного интеллекта Aidos-X, наиболее достоверной моделью оказалась модель INF3, точность модели составила 0,994.

АСК-анализ, использованный в данной работе, позволяет:

- сформировать обобщенные лингвистические образы классов (семантические ядра) на основе фрагментов или примеров, относящихся к ним текстов на любом языке;
- количественно сравнить лингвистический образ конкретного героя, или описание объекта, процесса с обобщенными лингвистическими образами групп (классов);
- сравнить обобщенные лингвистические образы классов друг с другом и создавать их кластеры и конструкты;
- исследовать моделируемую предметную область путем исследования лингвистической системно-когнитивной модели.

Все это можно делать для любого естественного или искусственного языка, или системы кодирования.

СПИСОК ЛИТЕРАТУРЫ

1. Луценко Е.В. Синтез адаптивных интеллектуальных измерительных систем с применением АСК-анализа и системы «Эйдос» и системная идентификация в эконометрике, биометрии, экологии, педагогике, психологии и медицине / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1 – 60. – IDA [article ID]: 1161602001. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/01.pdf>, 3,75 у.п.л.3.
2. Луценко Е.В. АСК-анализ, моделирование и идентификация живых существ на основе их фенотипических признаков / Е.В. Луценко, Ю.Н. Пенкина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: Куб- ГАУ, 2014. – №06(100). С. 1346 – 1395. – IDA [article ID]: 1001406090. – Режим доступа: <http://ej.kubagro.ru/2014/06/pdf/90.pdf>, 3,125 у.п.л.
3. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с.
4. Сайт профессора Е.В.Луценко [Электронный ресурс]. Режим доступа: <http://lc.kubagro.ru/>, свободный. - Загл. с экрана. Яз.рус.
5. Луценко Е.В. Количественная оценка степени манипулирования индексом Хирша и его модификация, устойчивая к манипулированию / Е.В. Луценко, А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №07(121). С. 202 – 234. – IDA [article ID]: 1211607005. – Режим доступа: <http://ej.kubagro.ru/2016/07/pdf/05.pdf>, 2,062 у.п.л.. <http://dx.doi.org/10.21515/1990-4665-121-0057>.
6. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос-Х++» Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2012. – №09(083). С. 328 – 356. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2012/09/pdf/25.pdf>, 1,812 у.п.л.
7. Луценко Е.В. Системная теория информации и нелокальные интерпретируемые нейронные сети прямого счета / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №01(001). С. 79 – 91. – IDA [article ID]: 0010301011. – Режим доступа: <http://ej.kubagro.ru/2003/01/pdf/11.pdf>, 0,812 у.п.л.
8. Луценко Е.В. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификации статей по научным специальностям с применением АСК-анализа и интеллектуальной системы «Эйдос» (на примере Научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии) / Е.В. Луценко, Н.В. Андрафанова, Н.В. Потапова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2019. – №01(145). С. 31 – 102. – IDA [article ID]: 1451901033. – Режим доступа: <http://ej.kubagro.ru/2019/01/pdf/33.pdf>, 4,5 у.п.л.