

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 лучших фильмов с 2020 по 2022 годы»

Выполнил студент группы: ИТ2141 Черкас Евгений Сергеевич

Допущен к защите _____

Руководитель проекта д.э.н., к.т.н., профессор Луценко Е. В. ()
(подпись, расшифровка подписи)

Защищен _____

Оценка _____

Краснодар

2023

РЕФЕРАТ

Курсовая работа содержит: 31 страниц, 31 рисунков, 10 литературных источников.

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

Цель данной курсовой работы заключается в создании интеллектуальных моделей, которые в полной мере отражают существующие причинно-следственные взаимосвязи популярности современных фильмов.

Для достижения поставленной цели применяется Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос».

В данной курсовой работе необходимо проанализировать методы формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами, принятия решений и исследования моделируемой предметной области путем исследования ее модели.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1 СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	5
1.1 Описание решения	5
1.2 Поиск и объединение исходных данных	7
1.3 Ввод выборки в систему Aidos-X	7
1.4 Синтез и верификация моделей	10
1.5 Виды моделей Aidos-X	11
1.6 Результаты верификации моделей	12
2 РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	17
2.1 Решение задач идентификации	17
2.2 SWOT и PEST матрицы и диаграммы	20
2.3 Исследование моделируемой предметной области путем исследования ее модели	21
2.3.1 Когнитивные диаграммы классов	22
2.3.2 Агломеративная когнитивная кластеризация классов	22
2.3.3 Когнитивные диаграммы значений факторов	24
2.3.4 Агломеративная когнитивная кластеризация значений факторов ..	24
2.3.5 Нелокальные нейроны и нелокальные нейронные сети	26
2.3.6 3D-интегральные когнитивные карты	27
ЗАКЛЮЧЕНИЕ	28
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ	29

ВВЕДЕНИЕ

Технологии искусственного интеллекта быстро развиваются и применяются в самых разных областях, таких как финансы, тяжелая промышленность, здравоохранение, управление персоналом, музыка, журналистика, обслуживание клиентов и многое другое. Искусственный интеллект используется в алгоритмической торговле, маркетинговых исследованиях, управлении финансовым портфелем, опасной работе с роботами, медицинской диагностике, уходе за пожилыми людьми и пациентами, подборе персонала и управлении им, автоматизированной журналистике, поддержке клиентов и во многих других отраслях.

Данная курсовая работа посвящена решению задачи анализа, а именно анализу популярности современных фильмов. Выявление зависимостей между наблюдаемыми данными и рейтингом позволит облегчить процесс предсказания будущих жанров и режиссеров фильмов.

Из всего вышесказанного можно выделить цель работы. Целью данной работы является разработка интеллектуальных моделей, отражающих реально существующие причинно-следственные взаимосвязи между некоторыми характеристиками фильмов и их популярностью.

Для достижения поставленной цели в данной работе будет применена технология автоматизированного системно-когнитивного анализа (АСК-анализ), который будет осуществлён с помощью интеллектуальной системы «AidosX++».

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.

Главным фактором выбора АСК-анализа является то, что он включает в себя не только теорию, но и метод количественного выявления в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных данных крайне большой размерности числовой и не числовой природы, которые являются измеряемыми в различных типах шкал и единицах измерения.

Универсальная когнитивная аналитическая система Aidos-X в настоящее время выступает очень важным, развитым и доступным программным инструментарием для реализации АСК-анализа.

Преимущества данной системы:

- универсальность;
- содержит большое количество локальных и облачных учебных и научных приложений (в настоящее время их 31 и 376, соответственно);
- мультиязычная поддержка интерфейса (больше 50 языков);
- поддерживает online-среду накопления знаний и широко используется во всем мире;
- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решения этих задач в несколько тысяч раз;
- доступность, то есть данная система находится в полном открытом бесплатном доступе, причем с актуальными исходными текстам;

- одна из первых отечественных систем искусственного интеллекта персонального уровня, т.е. она не требует от пользователя специальной подготовки в области технологий искусственного интеллекта (акт внедрения системы Aidos-X от 1987 года);
- 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;
- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение 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 (развитая 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.

В данной работе будет реализовано решение поставленной задачи в четыре этапа в соответствии с методологией АСК-анализа:

1. Поиск и объединение исходных данных в единую таблицу и сохранение в файле MS Excel.
2. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы Aidos-X.
3. Синтез и верификация моделей предметной области.
4. Применение моделей для решения задач идентификации, прогнозирования и исследования предметной области.

1.2 Поиск и объединение исходных данных

Входные данные, необходимые для обучения модели, были взяты из открытых источников. Итоговый файл содержит следующие данные:

- Title – название фильма;
- Genre – жанр фильма;
- Directors – режиссер фильма;
- Rating – рейтинг фильма;
- Year – год выпуска;
- Certificate – рейтинг;
- Runtime – продолжительность;
- Budget – бюджет;
- Country_of_origin – страна происхождения.

Фрагмент исходных данных в виде таблицы MS Excel представлен на рисунке 1.

	A	B	C	D	E	F	G	H	I	J	
1	Title	Genre	Directors	Rating	Year	Month	Certificate	Runtime	Budget	Income	Country_of_origin
2	Avatar: The Way of Water	Action, Adventure, Fantasy	James Cameron	7,8	2022	December	PG-13	192	\$350,000,000	\$2,267,946,983	United States
3	Guillermo del Toro's Pinocchio	Animation, Drama, Family	Guillermo del Toro, Mark Gustafson	7,6	2022	December	PG	117	\$35,000,000	\$108,967	United States, Mexico
4	Bullet Train	Action, Comedy, Thriller	David Leitch	7,3	2022	August	R	127	\$85,900,000	\$239,288,602	Japan, United States
5	The Banshees of Inisherin	Comedy, Drama	Martin McDonagh	7,8	2022	November	R	114	Unknown	\$19,720,823	Ireland, United Kingdom
6	M3gan	Horror, Sci-Fi, Thriller	Gerard Johnstone	6,4	2022	December	PG-13	102	\$12,000,000	\$171,253,910	United States
7	Emancipation	Action, Thriller	Antoine Fuqua	6,1	2022	December	R	132	\$120,000,000	Unknown	United States
8	Amsterdam	Comedy, Drama, History	David O Russell	6,1	2022	October	R	134	\$80,000,000	\$31,245,810	United States, Japan
9	Violent Night	Action, Comedy, Crime	Tommy Wirkola	6,9	2022	December	R	112	\$20,000,000	\$59,595,460	United States, Canada
10	The Whale	Drama	Darren Aronofsky	7,8	2022	December	R	117	\$10,000,000	\$31,104,312	United States
11	The Fabelmans	Drama	Steven Spielberg	7,6	2022	November	PG-13	151	\$40,000,000	\$9,500,361	United States
12	The Menu	Comedy, Horror, Thriller	Mark Mylod	7,5	2022	November	R	107	\$35,000,000	\$65,878,071	United States
13	Babylon	Comedy, Drama, History	Damien Chazelle	7,3	2022	December	R	188	\$78,000,000	\$61,851,455	United States
14	X	Horror, Mystery, Thriller	Ti West	6,6	2022	March	R	105	\$1,000,000	\$14,779,858	United States, Canada
15	Bones and All	Drama, Horror, Romance	Luca Guadagnino	7	2022	November	R	131	\$16,000,000	\$14,134,907	Italy, United States
16	Black Adam	Action, Adventure, Fantasy	Jaume Collet Serra	6,5	2022	October	PG-13	125	\$195,000,000	\$591,273,355	United States, Canada
17	Spirited	Comedy, Family, Musical	Sean Anders	6,6	2022	November	PG-13	127	\$100,000,000	Unknown	United States
18	Lady Chatterley's Lover	Drama, Romance	Laure de Clermont Tonnerre	6,7	2022	December	R	126	Unknown	Unknown	United Kingdom, United States
19	A Christmas Story Christmas	Comedy, Family	Clay Kaytis	6,8	2022	November	PG	98	Unknown	Unknown	United States, Mexico
20	Troll	Action, Adventure, Drama	Roar Uthaug	5,8	2022	December	TV-14	101	Unknown	Unknown	Norway
21	Black Panther: Wakanda Forever	Action, Adventure, Drama	Ryan Coogler	6,8	2022	November	PG-13	161	\$250,000,000	\$791,852,924	United States
22	Glass Onion: A Knives Out Mystery	Comedy, Crime, Drama	Rian Johnson	7,6	2022	December	PG-13	140	\$40,000,000	\$13,280,000	United States
23	Triangle of Sadness	Comedy, Drama	Ruben Ostlund	7,6	2022	September	R	147	€ 10,000,000	\$17,211,145	Sweden, France, United States
24	Everything Everywhere All at Once	Action, Adventure, Comedy	Dan Kwan, Daniel Scheinert	8,1	2022	March	R	139	\$25,000,000	\$103,384,361	United States
25	Emily the Criminal	Crime, Drama, Thriller	John Patton Ford	6,7	2022	August	R	97	\$3,000,000	\$2,156,296	United States
26	God's Crooked Lines	Mystery	Oriol Paulo	7	2022	December	TV-MA	154	Unknown	\$6,060,070	Spain
27	Don't Worry Darling	Drama, Thriller	Olivia Wilde	6,2	2022	September	R	123	\$20,000,000	\$66,709,493	United States
28	Pinocchio	Adventure, Comedy, Drama	Robert Zemeckis	5,1	2022	September	PG	105	\$150,000,000	\$37,253	United States
29	Top Gun: Maverick	Action, Drama	Joseph Kosinski	8,4	2022	May	PG-13	130	\$170,000,000	\$1,488,732,821	United States
30	Smile	Horror, Mystery, Thriller	Parker Finn	6,6	2022	September	R	115	\$17,000,000	\$216,135,048	United States
31	The Batman	Action, Crime, Drama	Matt Reeves	7,8	2022	March	PG-13	176	\$200,000,000	\$770,836,163	United States
32	Something from Tiffany's	Comedy, Family, Romance	Daryl Wein	6,2	2022	December	PG	87	Unknown	Unknown	United States
33	Barbarian	Horror, Mystery, Thriller	Zach Cregger	7,1	2022	September	R	102	\$4,500,000	\$45,352,337	United States
34	She Said	Drama, History	Maria Schrader	7,2	2022	November	R	129	\$32,000,000	\$11,110,294	United States
35	Ticket to Paradise	Comedy, Romance	Ol Parker	6,2	2022	October	PG-13	104	\$60,000,000	\$167,456,898	United States, United Kingdom
36	The Noel Diary	Comedy, Drama, Romance	Charles Shyer	6,1	2022	November	TV-14	99	Unknown	Unknown	United States
37	Prey for the Devil	Horror, Thriller	Daniel Stamm	5,2	2022	October	PG-13	93	Unknown	\$43,300,652	United States

Рисунок 1 – Фрагмент обучающей выборки

1.3 Ввод выборки в систему Aidos-X

Затем обучающая выборка в формате xls, импортирована в систему Aidos-X: скопирована в папку Inp_data и переименована в Inp_data.

Далее был использован универсальный программный интерфейс импорта данных в систему Aidos-X, а именно режим 2.3.2.2 (Рисунок 2).

В форме задания размерности модели системы указываем диапазоны столбцов и желаемое количество градаций описательных и классификационных шкал (Рисунок 3).

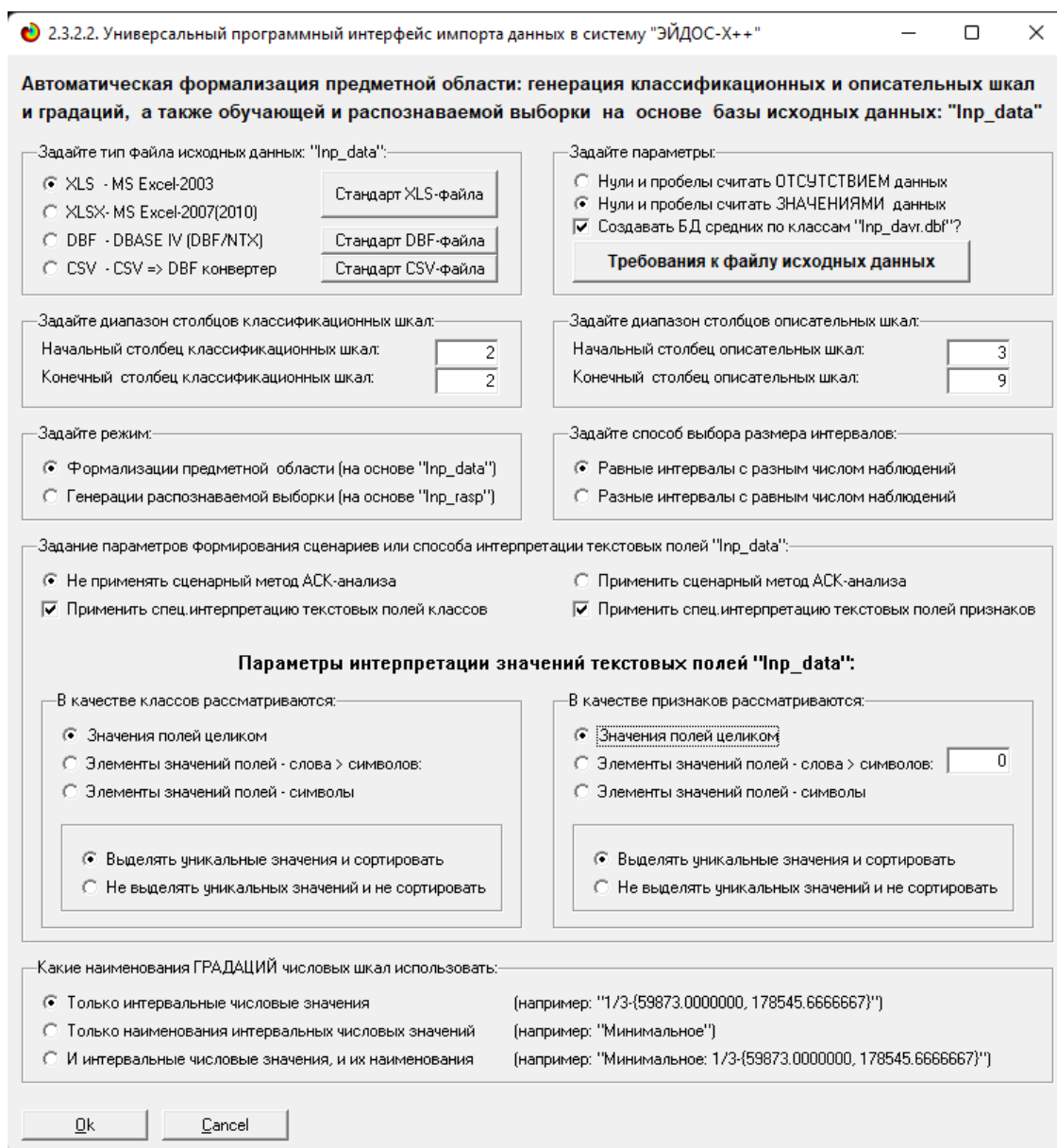


Рисунок 2 – Интерфейс импорта данных

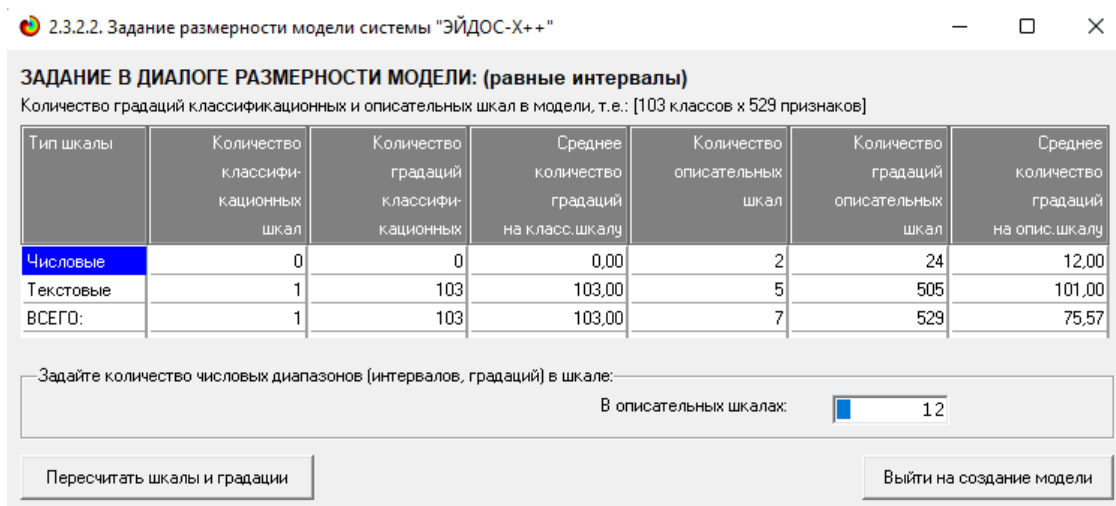


Рисунок 3 – Задание размерности модели

Процесс импорта данных из внешней БД "Inp_data.xls" в систему Aidos-X представлен на рисунке 4.

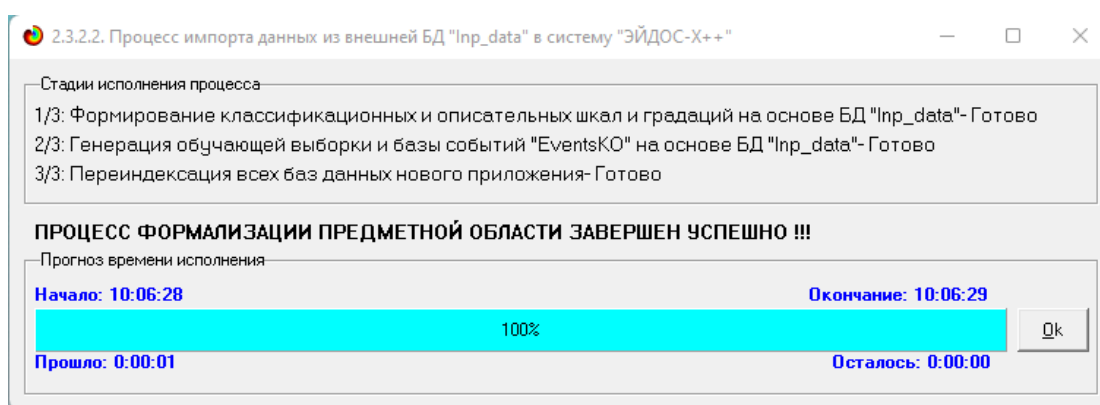


Рисунок 4 – Процесс импорта данных

Формирование классификационных и описательные шкал, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных, осуществляется автоматически после импорта. Тем самым этап формализации предметной области выполняется полностью автоматизировано.

1.4 Синтез и верификация моделей

Следующим шагом требуется запустить режим 3.5, в котором задаются модели для синтеза и верификации, а также задается модель, которой по окончании режима присваивается статус текущей (рисунок 5).

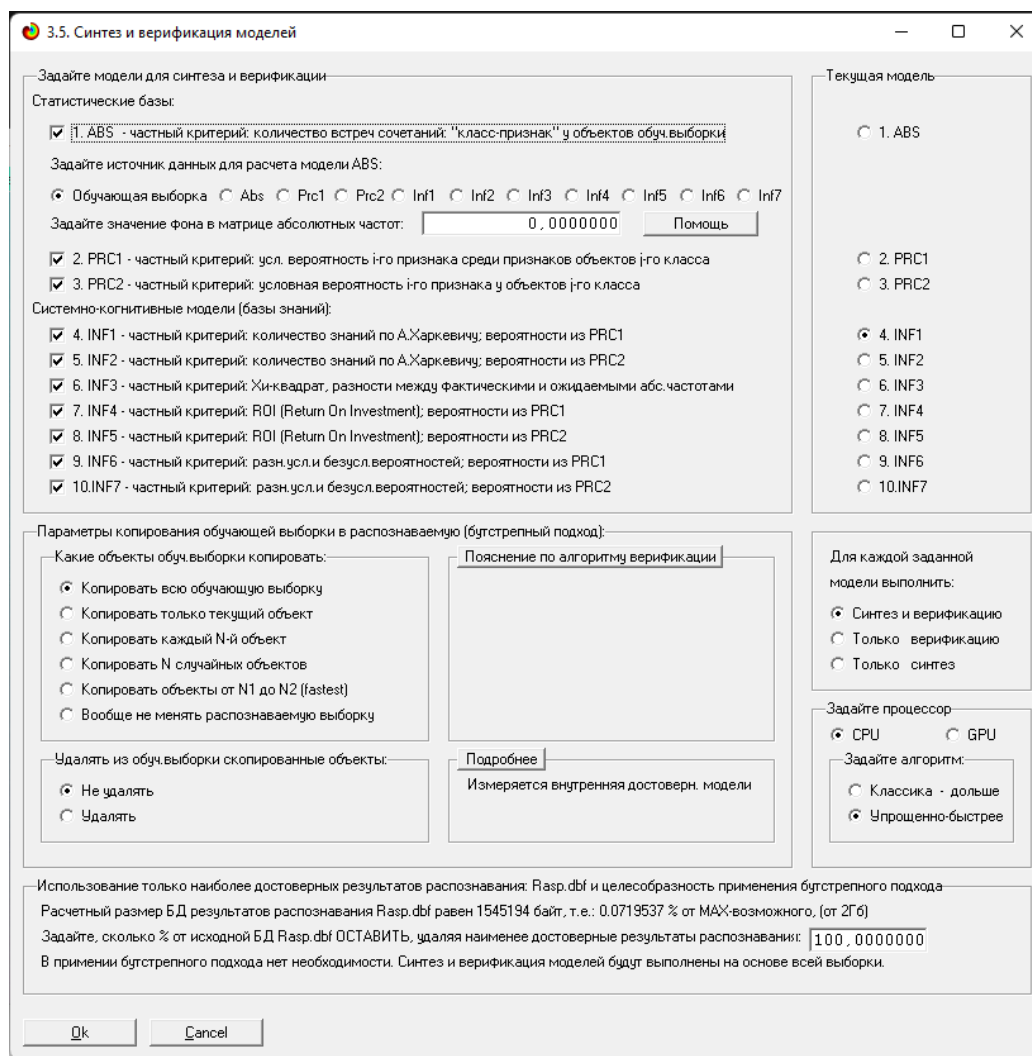


Рисунок 5 – Выбор моделей для синтеза и верификации

В данном режиме имеется много различных методов верификации моделей. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 6.

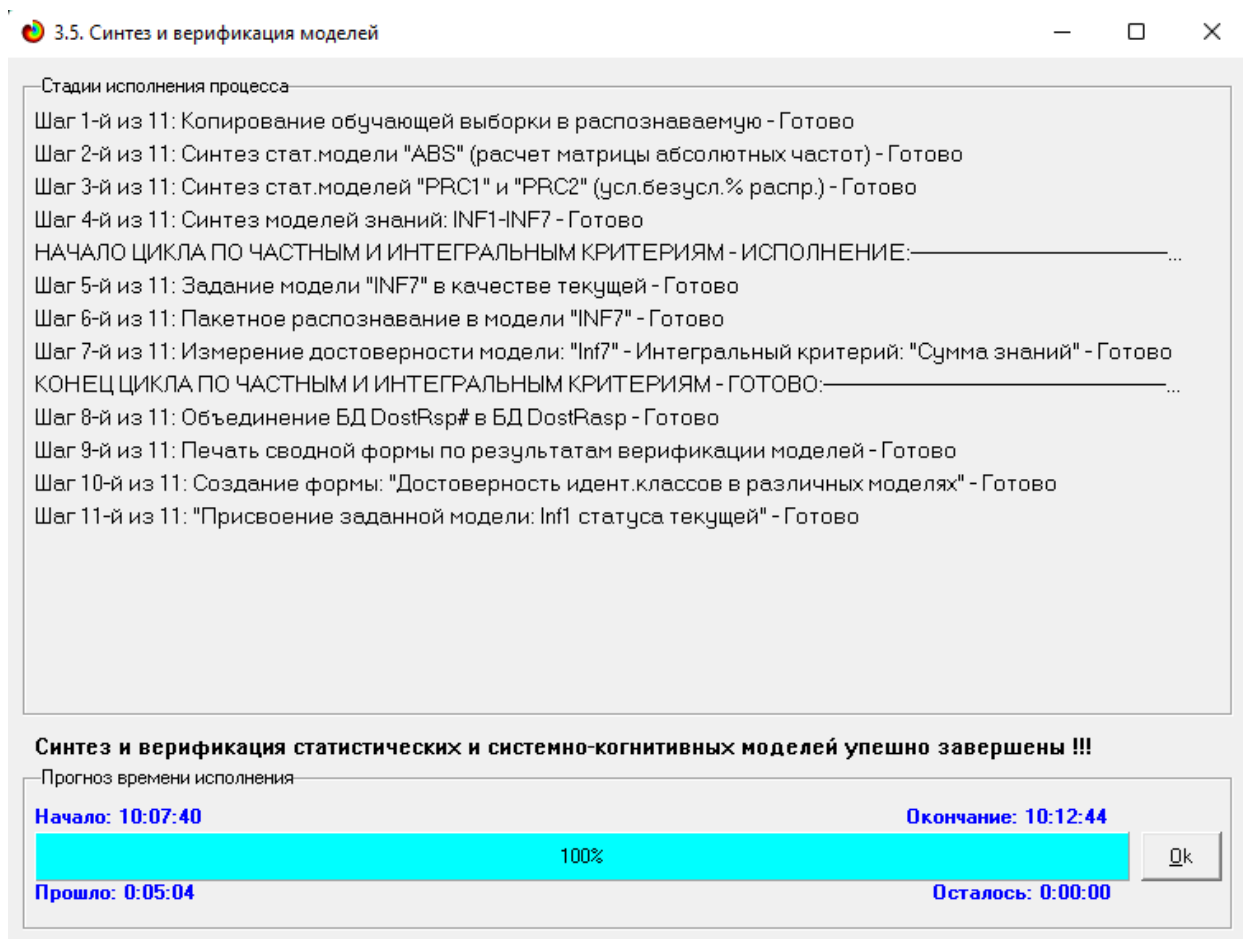


Рисунок 6 – Процесс исполнения режима 3.5

Стоит отметить, что синтез и верификация всех 10 моделей на данной задаче заняли 5 минут и 4 секунды. Далее перейдем непосредственно к выбору наиболее достоверной модели.

1.5 Виды моделей Aidos-X

На рисунках 7 и 8 приведены фрагменты созданных статистических и системно-когнитивных моделей (СК-модели).

Рассмотрим решение задачи идентификации на примере нескольких моделей, в которых рассчитано количество информации по А. Харкевичу, которое было получено по принадлежности идентифицируемого объекта к каждому из классов, если известно, что у этого объекта есть некоторый признак.

5.5. Модель: "1. ABS - частный критерий: количество встреч сочетаний "Класс-признак" у объектов обуч.выборки"

Код признака	Наименование описательной шкалы и градации	1. GENRE ACTION, ADVENTURE	2. GENRE ACTION, ADVENTURE, COMEDY	3. GENRE ACTION, ADVENTURE, CRIME	4. GENRE ACTION, ADVENTURE, DRAMA	5. GENRE ACTION, ADVENTURE, FANTASY	6. GENRE ACTION, ADVENTURE, SCI FI	7. GENRE ACTION, ADVENTURE, THRILLER
106.0	DIRECTORS-Hern n Jim nez							
107.0	DIRECTORS-Ilya Naishuller							
108.0	DIRECTORS-J Blakeson							
109.0	DIRECTORS-J D Dillard							
110.0	DIRECTORS-James Cameron					1.0		
111.0	DIRECTORS-James Gunn		1.0					
112.0	DIRECTORS-James Wan							
113.0	DIRECTORS-Jane Campion							
114.0	DIRECTORS-Janeen Damian							
115.0	DIRECTORS-Janicca Bravo							
116.0	DIRECTORS-Jared Bush, Byron Howard, Charise Castro Smith							
117.0	DIRECTORS-Jason Reitman							
118.0	DIRECTORS-Jason Volner							
119.0	DIRECTORS-Jaume Collet Serra		1.0			1.0		
120.0	DIRECTORS-Jeff Fowler		1.0					
121.0	DIRECTORS-Jeff W'adlow							
122.0	DIRECTORS-Jeremy Garelick							
123.0	DIRECTORS-Jerrod Carmichael							
124.0	DIRECTORS-Jeymes Samuel							
125.0	DIRECTORS-Jim Cummings							
126.0	DIRECTORS-Jim O Hanlon							
127.0	DIRECTORS-Joachim Trier							
128.0	DIRECTORS-Joe Begos							
129.0	DIRECTORS-Joe Camahan		1.0					

Рисунок 7 – Фрагмент матрицы абсолютных частот

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами"

Код признака	Наименование описательной шкалы и градации	1. GENRE ACTION, ADVENTURE	2. GENRE ACTION, ADVENTURE, COMEDY	3. GENRE ACTION, ADVENTURE, CRIME	4. GENRE ACTION, ADVENTURE, DRAMA	5. GENRE ACTION, ADVENTURE, FANTASY	6. GENRE ACTION, ADVENTURE, SCI FI
1.0	DIRECTORS-Aaron Schneider	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
2.0	DIRECTORS-Aaron Sorkin	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
3.0	DIRECTORS-Abhishek Pathak	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
4.0	DIRECTORS-Adam McKay	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
5.0	DIRECTORS-Adam Shankman	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
6.0	DIRECTORS-Adam Wingard	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
7.0	DIRECTORS-Adil El Abbi, Bilal Fallah	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
8.0	DIRECTORS-Alan Taylor	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
9.0	DIRECTORS-Alejandro G Imitu	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
10.0	DIRECTORS-Alex Ranarivelo	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
11.0	DIRECTORS-Alexon Locke	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
12.0	DIRECTORS-Anders Thomas Jensen	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
13.0	DIRECTORS-Andrew Dominik	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
14.0	DIRECTORS-Andy Seick	-0.003	-0.033	-0.010	-0.027	-0.030	0.09
15.0	DIRECTORS-Aneesh Chaganty	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
16.0	DIRECTORS-Angeliki Antoniou	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
17.0	DIRECTORS-Aniuth Iyer	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
18.0	DIRECTORS-Anthony Russo, Joe Russo	-0.007	-0.067	-0.020	-0.053	-0.060	-0.060
19.0	DIRECTORS-Antoine Fuqua	-0.010	-0.100	-0.030	-0.080	-0.090	-0.090
20.0	DIRECTORS-Antonio Campos	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
21.0	DIRECTORS-Anubhuti Kashyap	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
22.0	DIRECTORS-Autumn de Wilde	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
23.0	DIRECTORS-Barbara Bialowas, Tomasz Mandes	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030
24.0	DIRECTORS-Baz Luhrmann	-0.003	-0.033	-0.010	-0.027	-0.030	-0.030

Рисунок 8 – Фрагмент модели INF3

1.6 Результаты верификации моделей

Результаты верификации моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями, были отражены на рисунке 9.

3.4. Обобщ. форма по достов. моделям при разн. крит. Текущая модель: "INF1"

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	Помехи модели	Фигура Ван Ризбергера	Сумма модулей уровней сход. истинно-полож. решений (STP)	Сумма модулей уровней сход. истинно-отриц. решений (STN)	Сумма модулей уровней ложно-полож. решений	Сумма модулей уровней ложно-отриц. решений
1. ABS - частный критерий: количество встреч сонетной "Улас...	Корреляция абс. частот с обр...	300	300	3199	27401		0.011	1.000	0.021	203.441	44.899	7036	
1.ABS - частный критерий: количество встреч сонетной "Улас...	Сумма абс. частот по признак...	300	300	3199	27401		0.011	1.000	0.021	110.923		3021	
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Корреляция усл.отг. частот с о...	300	300	3199	27401		0.011	1.000	0.021	203.441	44.899	7036	
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Сумма усл.отг. частот по при...	300	300	3199	27401		0.011	1.000	0.021	147.765		5741	
3. PRC2 - частный критерий: условная вероятность иго признака.	Корреляция усл.отг. частот с о...	300	300	3199	27401		0.011	1.000	0.021	203.440	44.899	7036	
3. PRC2 - частный критерий: условная вероятность иго признака.	Сумма усл.отг. частот по при...	300	300	3199	27401		0.011	1.000	0.021	147.765		5741	
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	300	300	10026	20574		0.014	1.000	0.029	134.630	235.391	1242	
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Сумма знаний	300	300	6646	23954		0.012	1.000	0.024	138.437	86.721	1346	
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	300	300	10026	20574		0.014	1.000	0.029	134.630	235.391	1242	
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Сумма знаний	300	300	6646	23954		0.012	1.000	0.024	138.437	86.721	1346	
6. INF3 - частный критерий: Уинкватт, разности между фактич.	Семантический резонанс зна...	300	300	16624	13976		0.021	1.000	0.041	168.833	2088.983	1880	
6. INF3 - частный критерий: Уинкватт, разности между фактич.	Сумма знаний	300	300	16624	13976		0.021	1.000	0.041	102.485	1018.445	915	
7. INF4 - частный критерий: ROI (Return On Investment), веротно.	Семантический резонанс зна...	300	300	27131	3469		0.080	1.000	0.147	117.245	336.606	69	
7. INF4 - частный критерий: ROI (Return On Investment), веротно.	Сумма знаний	300	300	5710	24890		0.012	1.000	0.024	60.919	1.438	109	
8. INF5 - частный критерий: ROI (Return On Investment), веротно.	Семантический резонанс зна...	300	300	27131	3469		0.080	1.000	0.147	117.245	336.606	69	
8. INF5 - частный критерий: ROI (Return On Investment), веротно.	Сумма знаний	300	300	5710	24890		0.012	1.000	0.024	60.919	1.438	109	
9. INF6 - частный критерий: разн.усли безул.вероятностей, вер...	Семантический резонанс зна...	300	300	7233	23367		0.013	1.000	0.025	167.116	369.650	3560	
9. INF6 - частный критерий: разн.усли безул.вероятностей, вер...	Сумма знаний	300	300	6317	24283		0.012	1.000	0.024	100.396	98.344	2764	
10.INF7 - частный критерий: разн.усли безул.вероятностей, ве...	Семантический резонанс зна...	300	300	7233	23367		0.013	1.000	0.025	167.116	369.650	3560	
10.INF7 - частный критерий: разн.усли безул.вероятностей, ве...	Сумма знаний	300	300	6326	24274		0.012	1.000	0.024	100.396	98.344	2764	

Рисунок 9 – Оценка достоверности моделей (F-критерий)

3.4. Обобщ. форма по достов. моделям при разн. крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	L1-мера проф. Е.В. Луценко	Средний модуль уровней сход. истинно-полож. решений	Средний модуль уровней сход. истинно-отриц. решений	Средний модуль уровней сход. ложно-полож. решений	Средний модуль уровней сход. ложно-отриц. решений	A-Точность модели AР=СТП/(АТР+АТР+)	A-Помехи модели АР=СТН/(АТР+АТР+)	L2-мера проф. Е.В. Луценко	Процент правильной идентификац.	Процент правильной не идентификац.	Процент ошиб. иденц.
1. ABS - частный критерий: количество встреч сонетной "Улас...	Корреляция абс. частот с обр...	0.055	0.678	0.014	0.257		0.725	1.000	0.841	100.000	4.243	5
1.ABS - частный критерий: количество встреч сонетной "Улас...	Сумма абс. частот по признак...	0.068	0.370		0.110		0.770	1.000	0.870	100.000	4.243	5
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Корреляция усл.отг. частот с о...	0.055	0.678	0.014	0.257		0.725	1.000	0.841	100.000	4.243	5
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Сумма усл.отг. частот по при...	0.049	0.493		0.210		0.702	1.000	0.825	100.000	4.243	5
3. PRC2 - частный критерий: условная вероятность иго признака.	Корреляция усл.отг. частот с о...	0.055	0.678	0.014	0.257		0.725	1.000	0.841	100.000	4.243	5
3. PRC2 - частный критерий: условная вероятность иго признака.	Сумма усл.отг. частот по при...	0.049	0.493		0.210		0.702	1.000	0.825	100.000	4.243	5
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	0.178	0.449	0.023	0.060		0.881	1.000	0.937	100.000	49.395	5
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Сумма знаний	0.171	0.461	0.013	0.056		0.891	1.000	0.943	100.000	28.676	5
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	0.178	0.449	0.023	0.060		0.881	1.000	0.937	100.000	49.395	5
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Сумма знаний	0.171	0.461	0.013	0.056		0.891	1.000	0.943	100.000	28.676	5
6. INF3 - частный критерий: Уинкватт, разности между фактич.	Семантический резонанс зна...	0.182	0.563	0.126	0.135		0.807	1.000	0.893	100.000	52.618	5
6. INF3 - частный критерий: Уинкватт, разности между фактич.	Сумма знаний	0.183	0.342	0.061	0.066		0.839	1.000	0.912	100.000	52.618	5
7. INF4 - частный критерий: ROI (Return On Investment), веротно.	Семантический резонанс зна...	0.774	0.391	0.012	0.020		0.952	1.000	0.975	100.000	90.648	5
7. INF4 - частный критерий: ROI (Return On Investment), веротно.	Сумма знаний	0.530	0.203	0.000	0.004		0.979	1.000	0.989	100.000	21.713	5
8. INF5 - частный критерий: ROI (Return On Investment), веротно.	Семантический резонанс зна...	0.774	0.391	0.012	0.020		0.952	1.000	0.975	100.000	90.648	5
8. INF5 - частный критерий: ROI (Return On Investment), веротно.	Сумма знаний	0.530	0.203	0.000	0.004		0.979	1.000	0.989	100.000	21.713	5
9. INF6 - частный критерий: разн.усли безул.вероятностей, вер...	Семантический резонанс зна...	0.086	0.557	0.051	0.152		0.785	1.000	0.880	100.000	30.479	5
9. INF6 - частный критерий: разн.усли безул.вероятностей, вер...	Сумма знаний	0.068	0.335	0.016	0.114		0.746	1.000	0.855	100.000	26.340	5
10.INF7 - частный критерий: разн.усли безул.вероятностей, ве...	Семантический резонанс зна...	0.086	0.557	0.051	0.152		0.785	1.000	0.880	100.000	30.479	5
10.INF7 - частный критерий: разн.усли безул.вероятностей, ве...	Сумма знаний	0.068	0.335	0.016	0.114		0.746	1.000	0.855	100.000	26.387	5

Рисунок 10 – Оценка достоверности моделей (L1 и L2 критерии)

Из рисунка 9-10 можно сделать вывод о том, что в данном интеллектуальном приложении по F-критерию Ван Ризбергера наиболее достоверными моделями являются INF4 и INF5 с интегральным критерием «Семантический резонанс знаний» ($F=0,816$ в максимуме 1,000), по критерию L1 профессора Е.В. Луценко наиболее достоверной является СК-модель INF4 и INF5 с интегральным критерием «Семантический резонанс знаний» ($L1=0,890$ при максимуме 1,000), а по критерию L2 профессора Е.В. Луценко наиболее достоверной является СК-модель INF4 и INF5 с интегральным критерием «Сумма знаний» ($L2=0,836$ при максимуме 1,000). Для оценки достоверности моделей в АСК- анализе и системе Aidos-X используется F-

мера Ван Ризбергера и L-мера, представляющая собой ее нечеткое мультиклассовое обобщение, предложенное профессором Е.В. Луценко (рисунок 11).

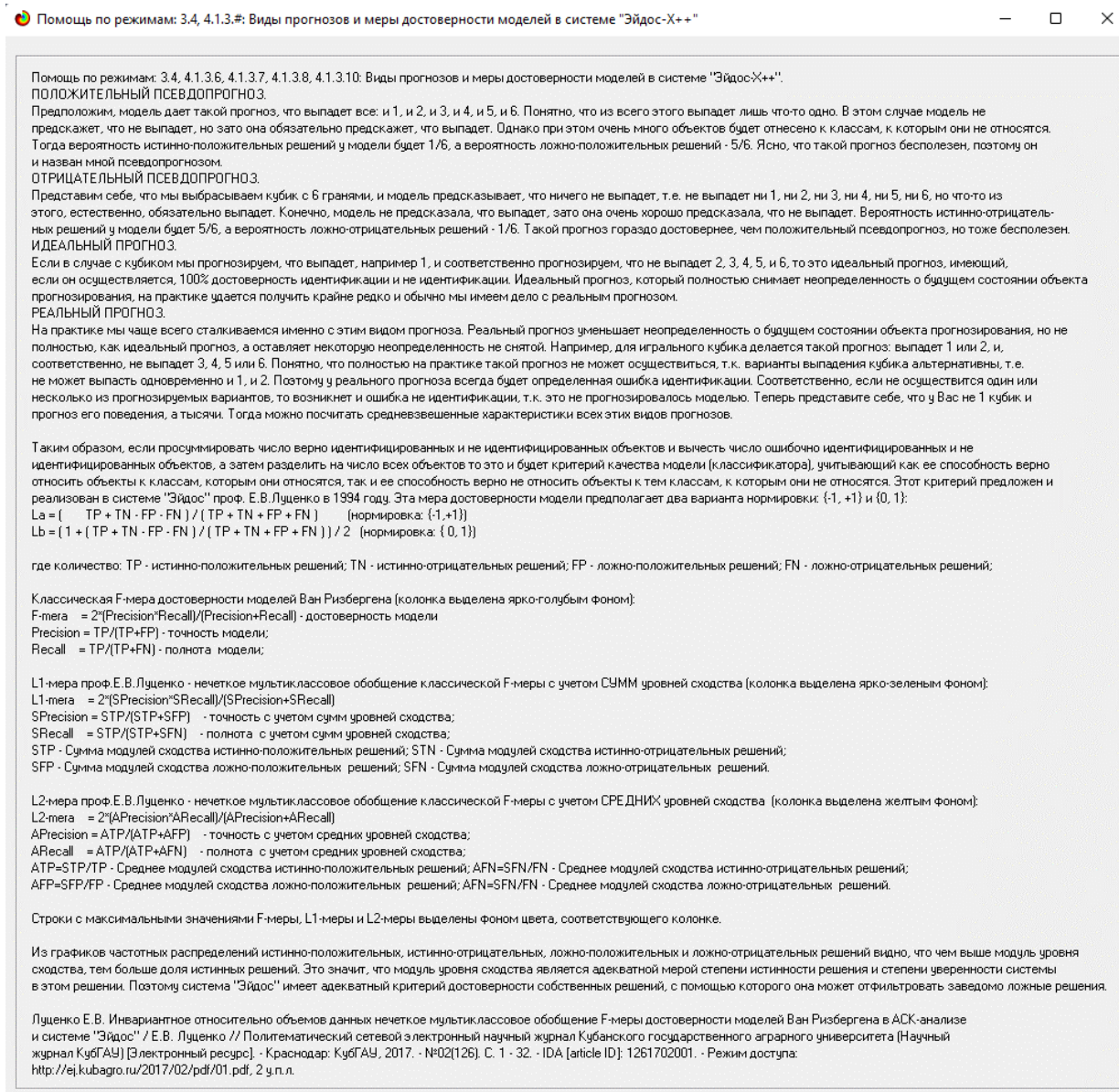


Рисунок 11 – Виды прогнозов и принципы определения

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.

На рисунках 12-15 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений в СК-моделях по данным обучающей выборки.

Левое распределение включает истинно-отрицательные и ложноположительные решения, а правое включает ложные отрицательные и истинно-положительные решения.

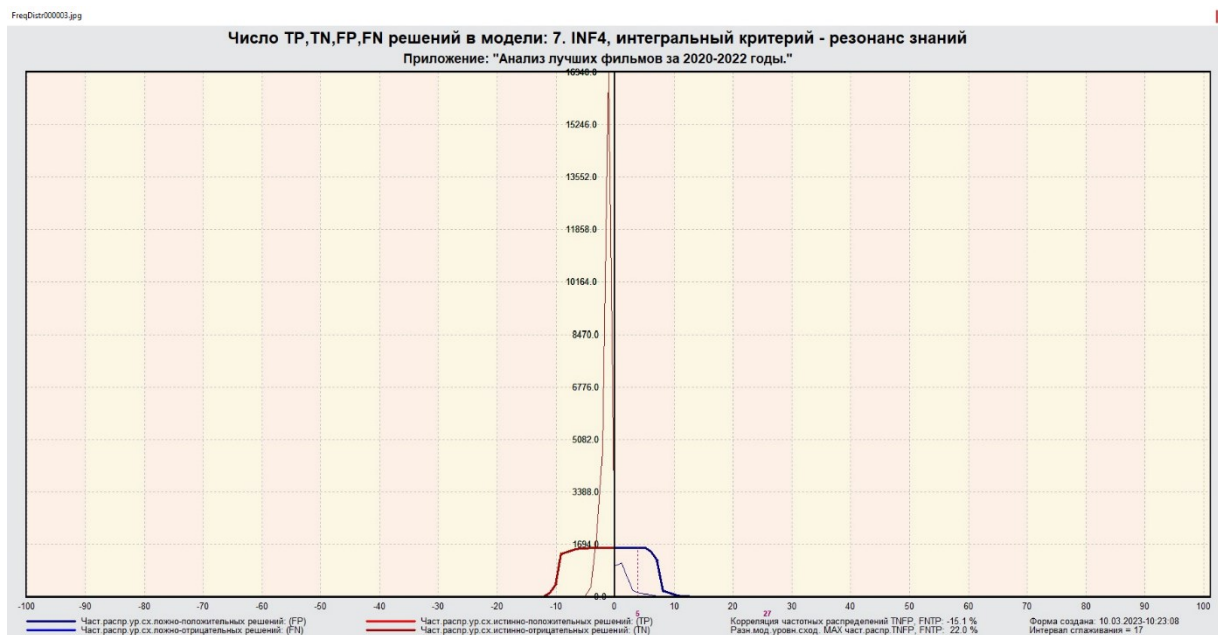


Рисунок 12 – Частотные распределения модели INF4 (резонанс знаний)

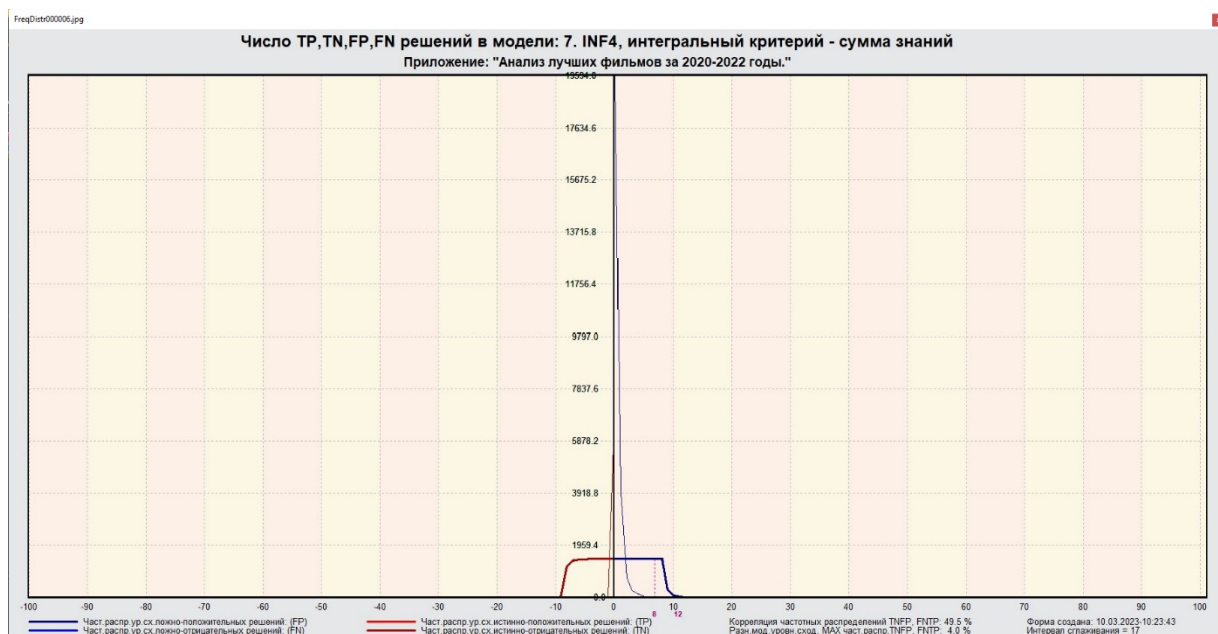


Рисунок 13 – Частотные распределения модели INF4 (сумма знаний)

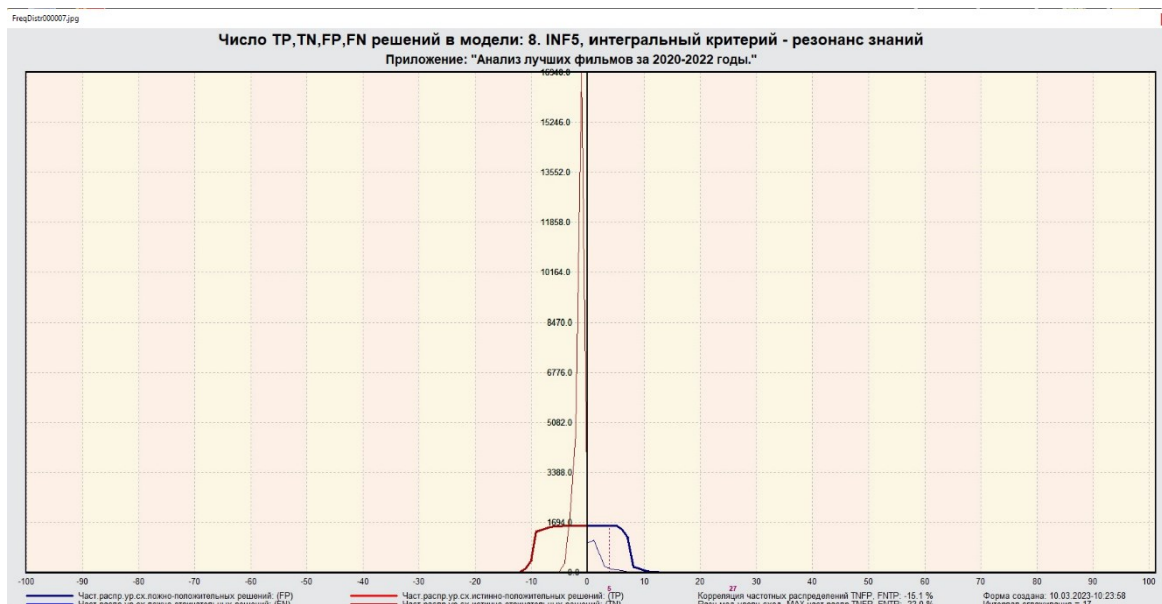


Рисунок 14 – Частотные распределения модели INF5 (резонанс знаний)

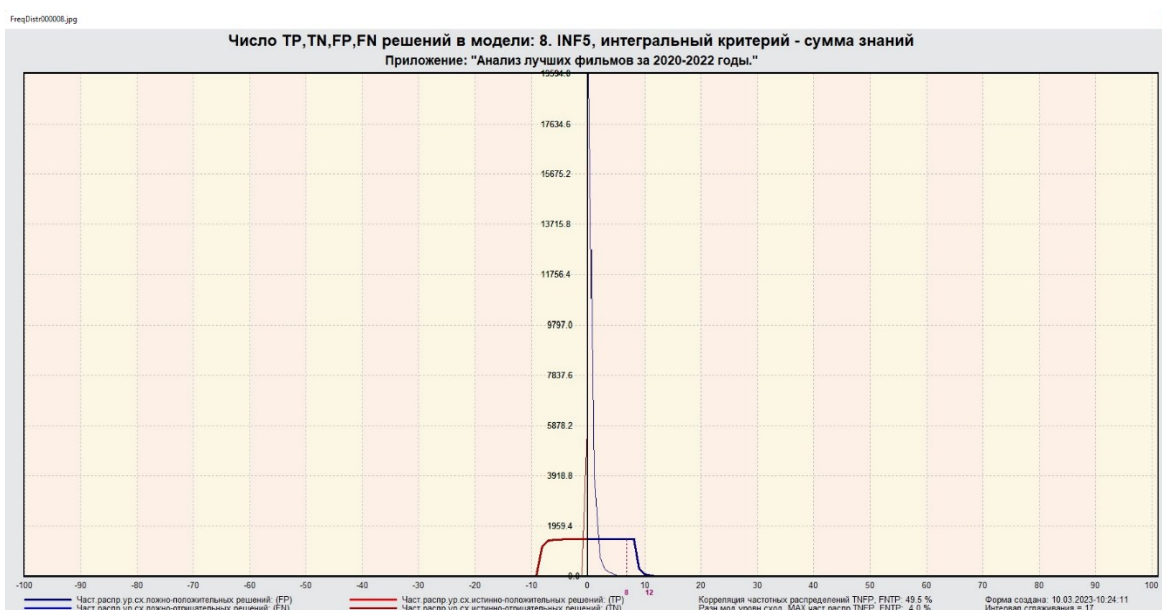


Рисунок 15 – Частотные распределения модели INF5 (сумма знаний)

Видно, что модели INF4 и INF5 на основе интегрального критерия «резонанс знаний» при проценте схожести более 70% практически не дают ложноотрицательных решений. При этом модели INF4 и INF5 на основе интегрального критерия «сумма знаний» при проценте схожести меньше примерно 15% дают меньше ложноположительных, чем истинно-положительных решений. Из этого можно сделать вывод о том, что система лучше определяет принадлежность объекта к классу, чем непринадлежность.

2 РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1 Решение задач идентификации

Решим задачу системной идентификации, т.е. определим к каким респонденты к каким категориям относятся на основе обучающей выборки в наиболее достоверной СК-модели INF4 (режим 5.6) (рисунок 16, 17).

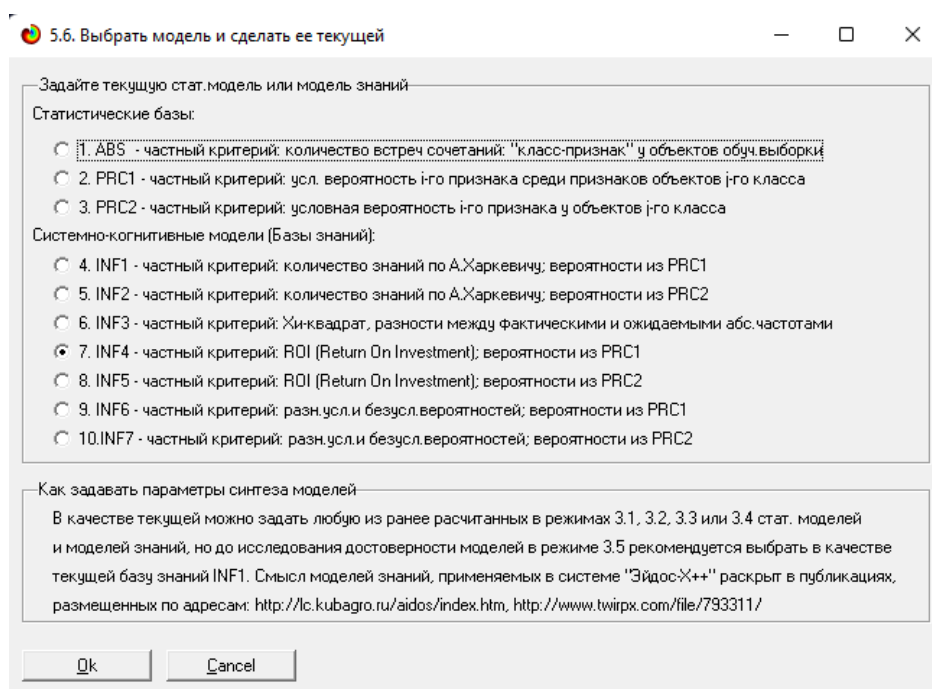


Рисунок 16 – Интерфейс задания модели

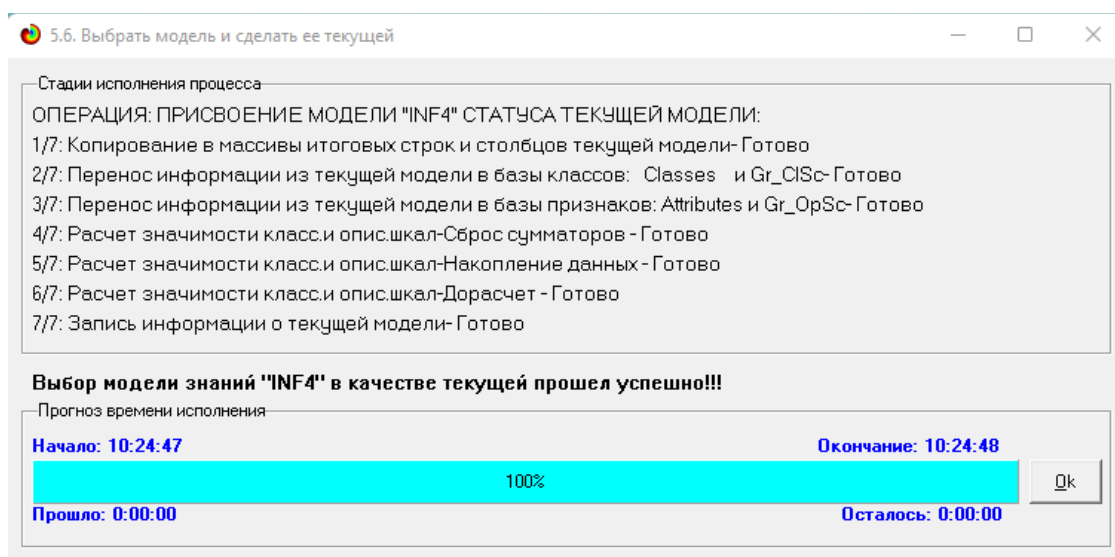


Рисунок 17 – Процесс загрузки модели

Затем следует провести пакетное распознавание в режиме 4.1.2. (рисунок 18).

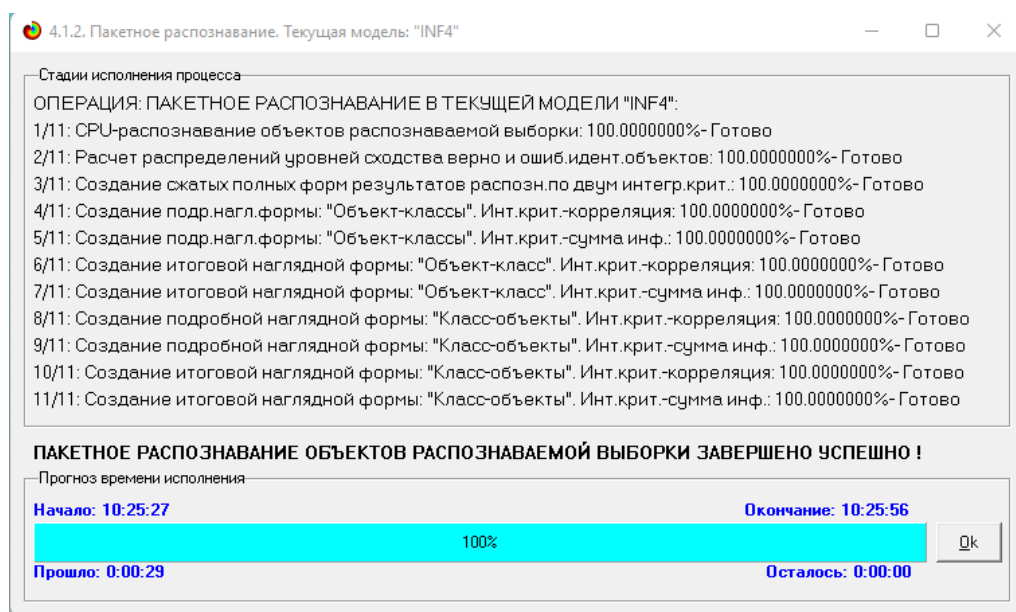


Рисунок 18 – Процесс пакетного распознавания модели

В результате пакетного распознавания в текущей модели создается ряд баз данных, которые визуализируются в выходных экранных формах, отражающих результаты решения задачи идентификации и прогнозирования.

Режим 4.1.3 системы Aidos-X 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.

На рисунке 19 и 20 приведен пример прогноза высокой и низкой достоверности частоты объектов и классов в модели INF4 на основе наблюдения предыстории их развития. Визуализация результатов распознавания представлена в соотношении: «Объект-классы» и «Класс-объекты».

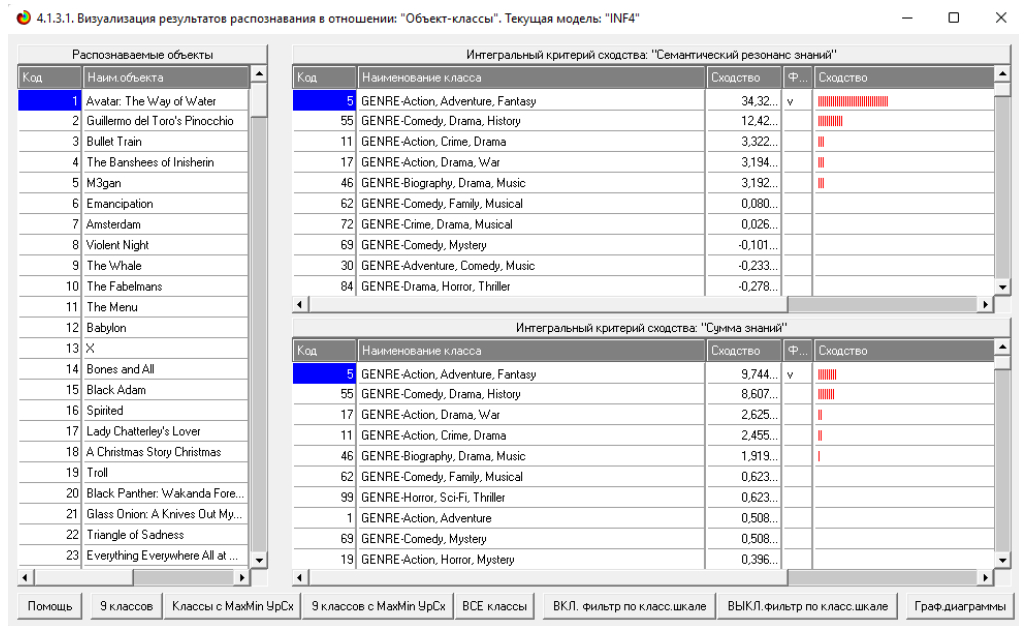


Рисунок 19 – Визуализация «Объект-классы»

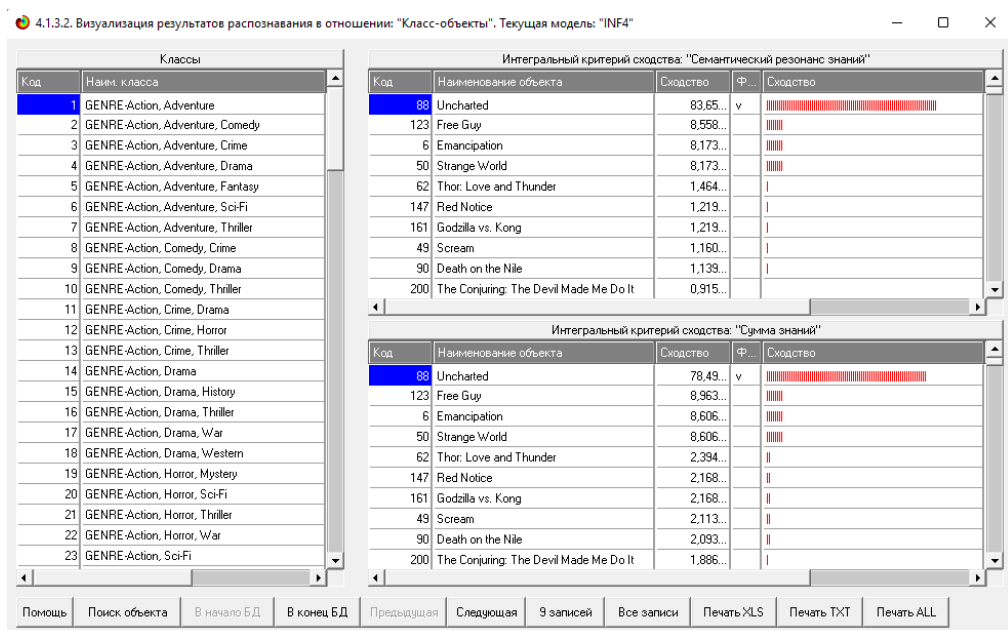


Рисунок 20 – Визуализация «Класс-объекты»

2.2 SWOT и PEST матрицы и диаграммы

При принятии решений определяется сила и направление влияния значений факторов на принадлежность состояний объекта моделирования к тем или иным классам, соответствующим различным будущим состояниям. По сути, это решение задачи SWOT-анализа.

Применительно к решаемой задаче, SWOT-анализ показывает, как различные критерии влияют на уровень заработной платы (рисунок 21).

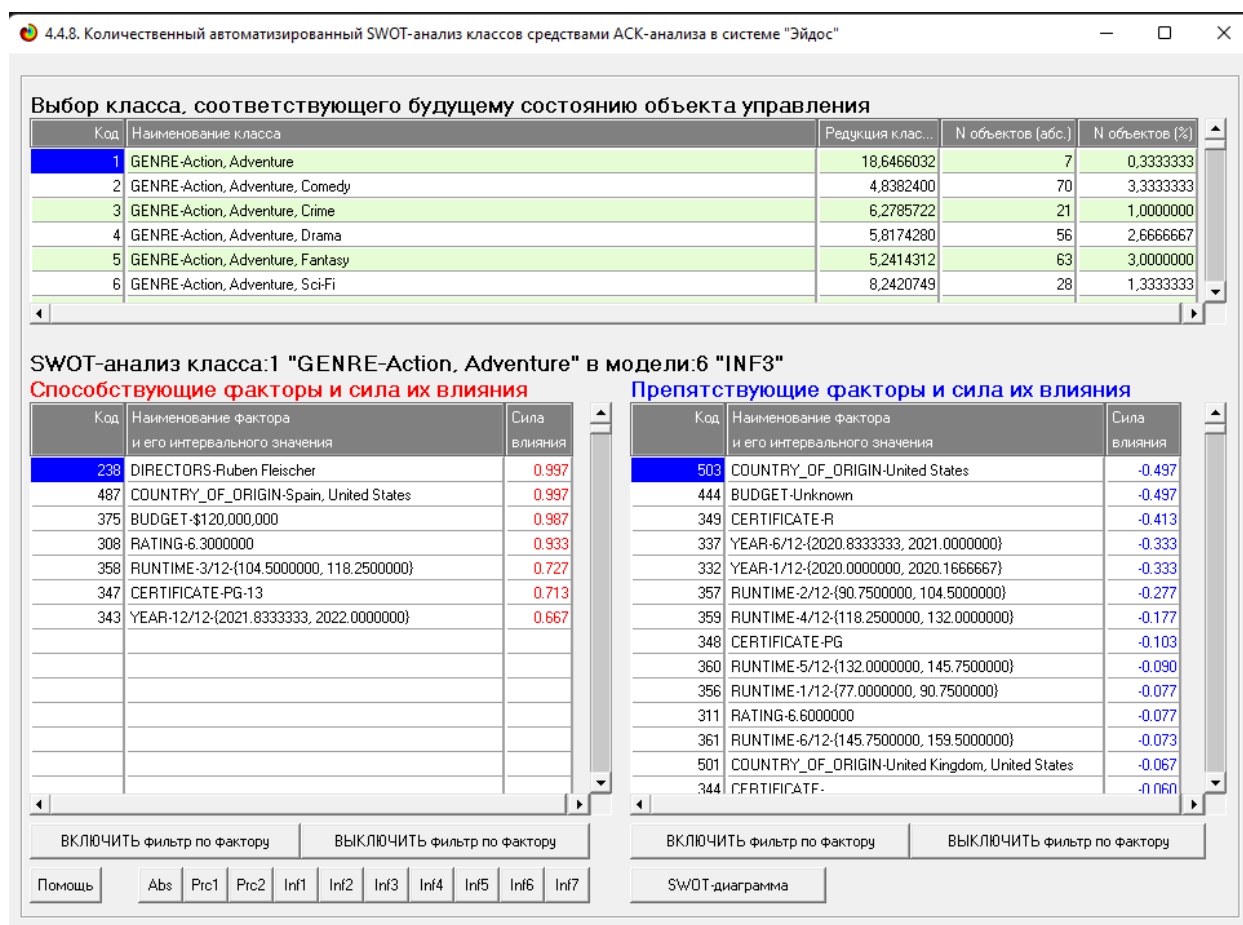


Рисунок 21 – SWOT-матрица

На рисунке 22 приведен пример табличной выходной формы количественного автоматизированного SWOT-анализа и PEST-анализа средствами системы Aidos-X.

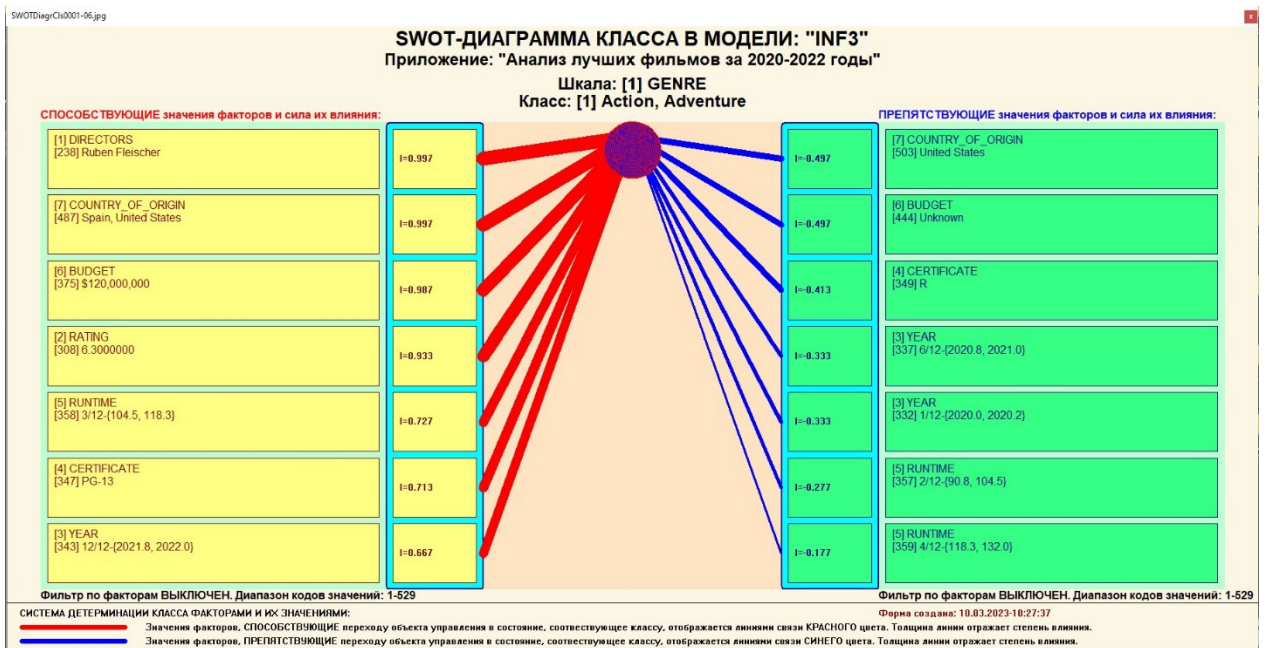


Рисунок 22 – SWOT-диаграмма

В левой стороне представлены факторы характерные для класса, а справа – нехарактерные.

2.3 Исследование моделируемой предметной области путем исследования ее модели

Если модель предметной области достоверна, то исследование модели можно считать исследованием самого моделируемого объекта, т.е. результаты исследования модели корректно относить к самому объекту моделирования, «переносить на него».

В системе «Aidos-X» есть довольно много возможностей для такого исследования, но в данной работе из-за ограничений на ее объем мы рассмотрим лишь результаты кластерно-конструктивного анализа классов и признаков (когнитивные диаграммы и дендрограммы), а также нелокальные нейроны, нелокальные нейронные сети, 3d-интегральные когнитивные карты и когнитивные функции.

2.3.1 Когнитивные диаграммы классов

Эти диаграммы отражают сходство/различие классов. Мы получаем их в режимах 4.2.2.1 и 4.2.2.2.

Отметим также, что на когнитивной диаграмме, приведенной на рисунке 23, показаны количественные оценки сходства/различия показателей. Важно, что эти результаты сравнения получены с применением системно-когнитивной модели, созданной непосредственно на основе эмпирических данных, а не как традиционно делается на основе экспертных оценок неформализуемым путем на основе опыта, интуиции и профессиональной компетенции.

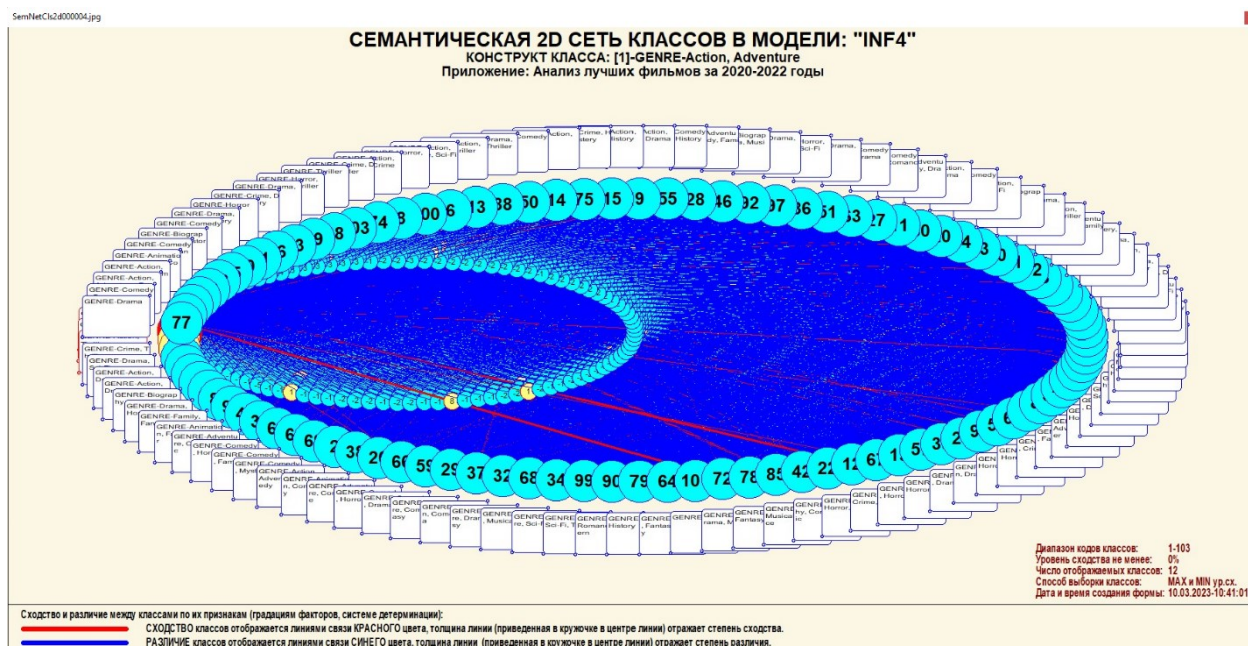


Рисунок 23 – Когнитивная диаграмма классов

2.3.2 Агломеративная когнитивная кластеризация классов

Информация о сходстве/различии классов, содержащаяся в матрице сходства, может быть визуализирована не только в форме, когнитивных диаграмм, но и в форме агломеративных дендрограмм с графиком изменения межкластерного расстояния, полученных в результате когнитивной кластеризации (рисунок 24 и 25).

ДЕНДРОГРАММА КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ КЛАССОВ В МОДЕЛИ: "INF4"
 Приложение: "Анализ лучших фильмов за 2020-2022 годы"

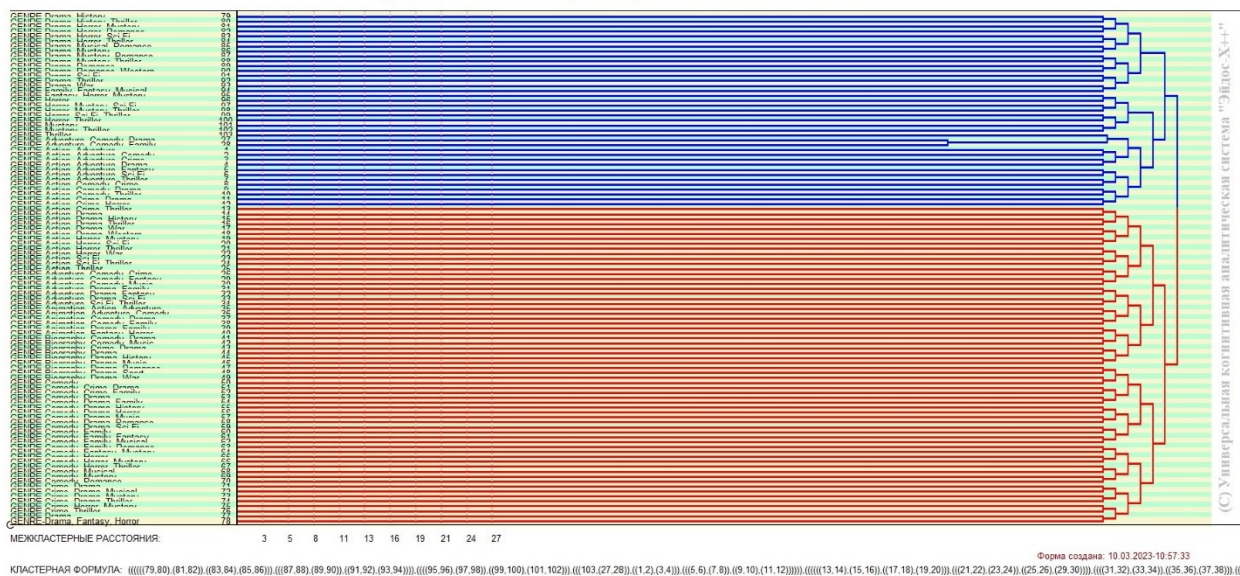


Рисунок 24 – Дендрограмма когнитивной кластеризации классов

ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ КЛАССОВ В МОДЕЛИ: "INF4"
 Приложение: "Анализ лучших фильмов за 2020-2022 годы"

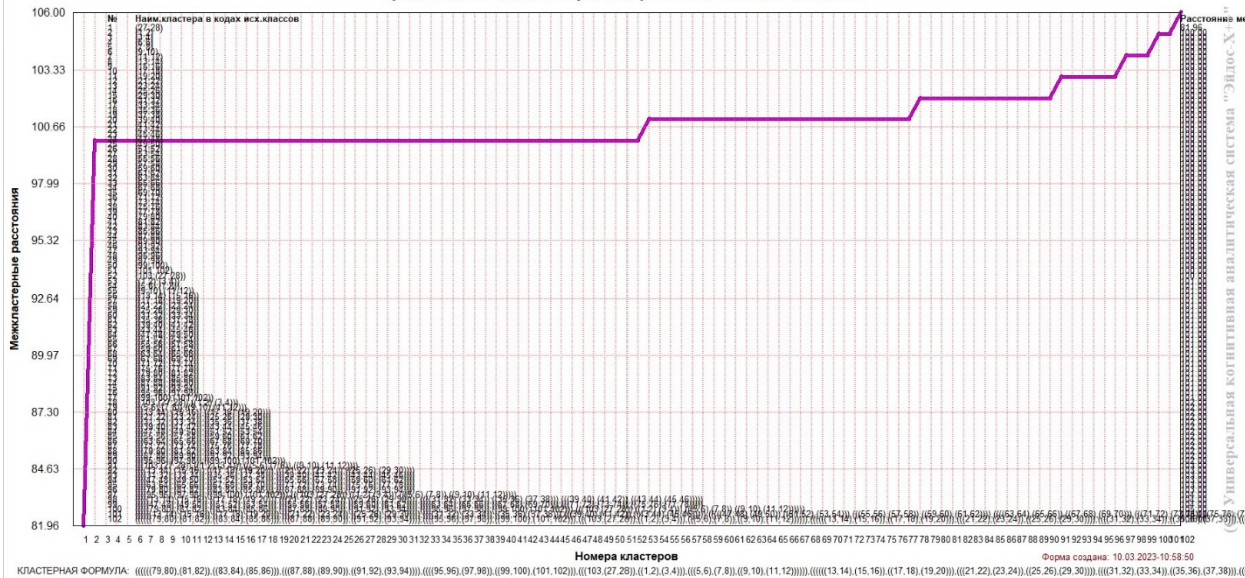


Рисунок 25 – График изменения межкластерных расстояний при когнитивной кластеризации классов

Из рисунка 24 мы видим, что некоторые характеристики сходны по системе детерминирующих их факторов, и, следовательно, могут наблюдаться одновременно, а другие по этой системе обуславливающих

факторов сильно отличаются, и, следовательно, являются взаимоисключающими, т.е. являются альтернативными.

2.3.3 Когнитивные диаграммы значений факторов

Подобные диаграммы отражают зависимость параметров. Эти диаграммы мы получаем в режимах 4.3.2.1 и 4.3.2.2 (рисунок 26).

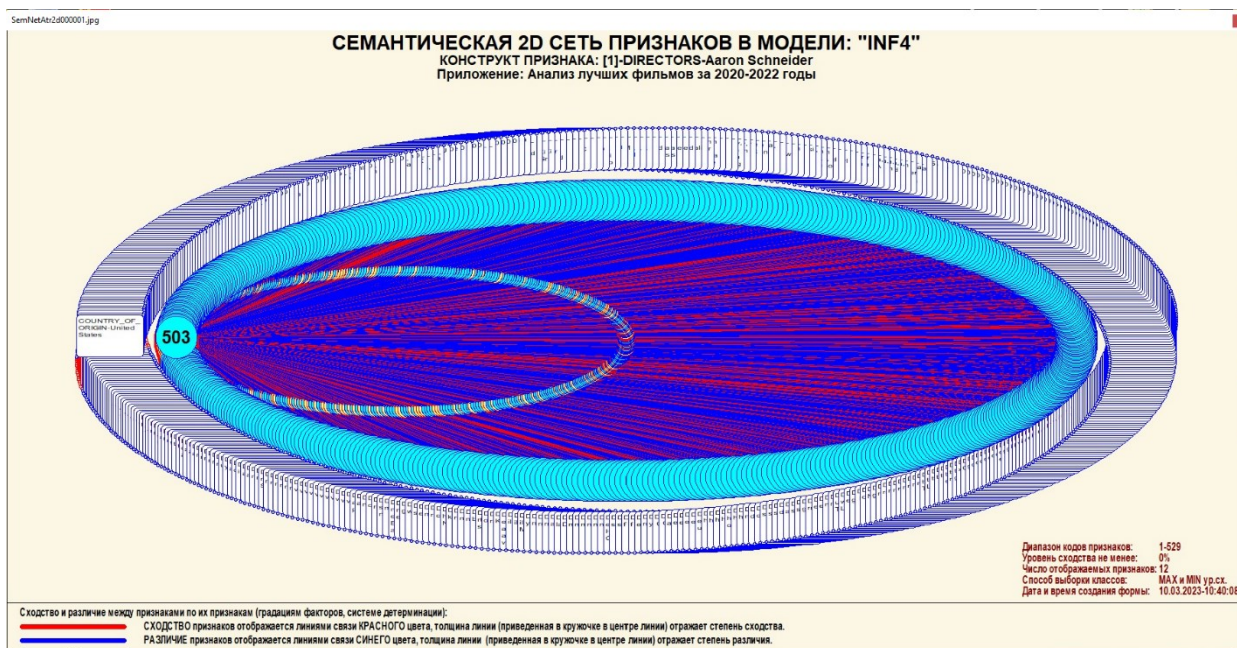


Рисунок 26 – График изменения межкластерных расстояний

2.3.4 Агломеративная когнитивная кластеризация значений факторов

На рисунках 27-28 приведена агломеративная дендрограмма когнитивной кластеризации значений факторов и график изменения межкластерных расстояний, полученные на основе той же матрицы сходства признаков по их смыслу, что и в когнитивных диаграммах.

ДЕНДРОГРАММА КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF4"
 Приложение: "Анализ лучших фильмов за 2020-2022 годы"

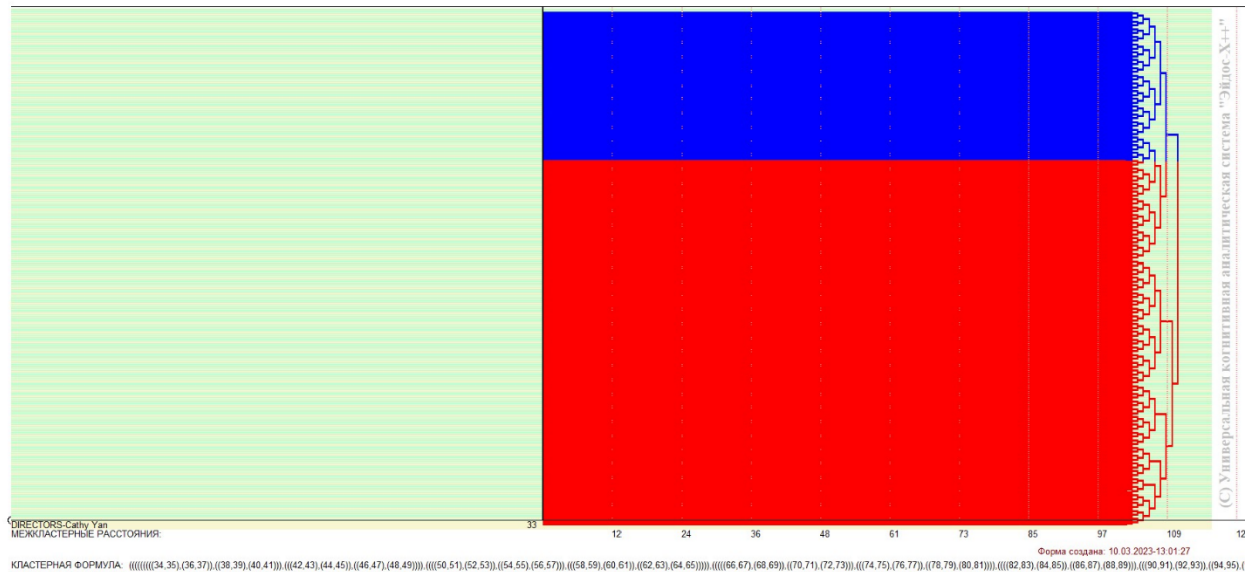


Рисунок 27 – Дендрограмма когнитивной кластеризации признаков

ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF4"
 Приложение: "Анализ лучших фильмов за 2020-2022 годы"

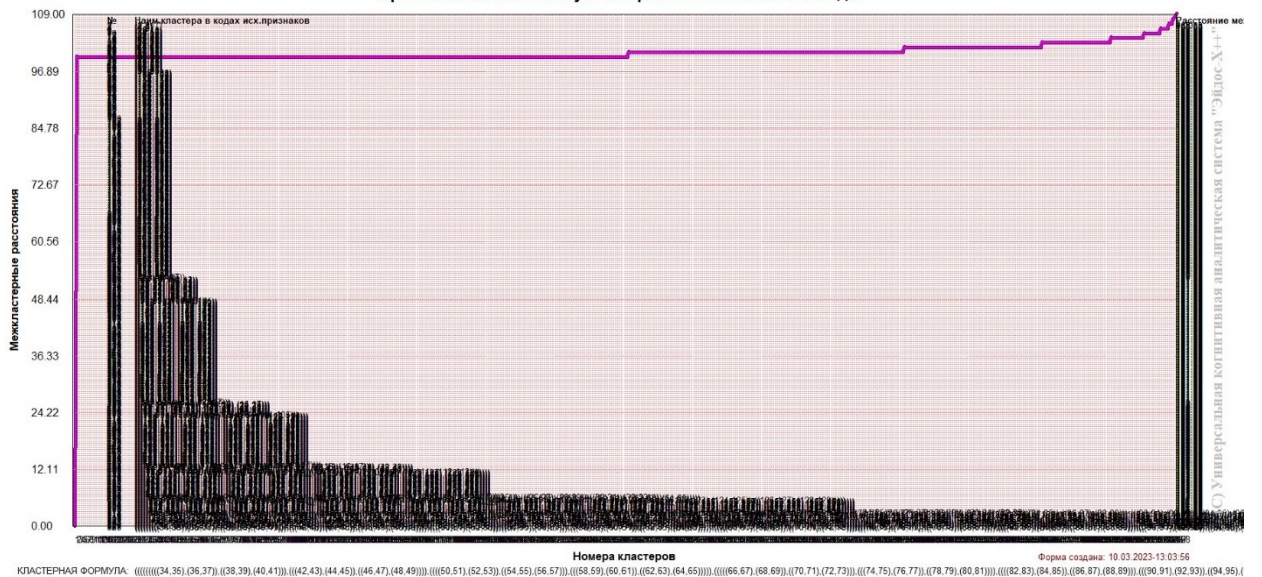


Рисунок 28 – График изменения межкластерных расстояний при когнитивной кластеризации признаков

Из дендрограммы мы видим, что все значения факторов образуют 2 четко выраженных кластера, объединенных в полюса конструкта (показаны синими и красным цветами). Хорошо видна группировка признаков по их смыслу.

2.3.5 Нелокальные нейроны и нелокальные нейронные сети

На рисунке 29 приведён пример нелокального нейрона, а на рисунке 30 – фрагмент одного слоя нелокальной нейронной сети, отражающий 87% наиболее значимых синаптических связей.

В приведенном фрагменте слоя нейронной сети нейроны соответствуют уровню заработной платы, а рецепторы – параметрам, влияющим на ее размер. Нейроны расположены слева направо в порядке убывания силы детерминации, т.е. слева находятся результаты, наиболее жестко обусловленные обуславливающими их значениями факторами, а справа – менее жестко обусловленные.

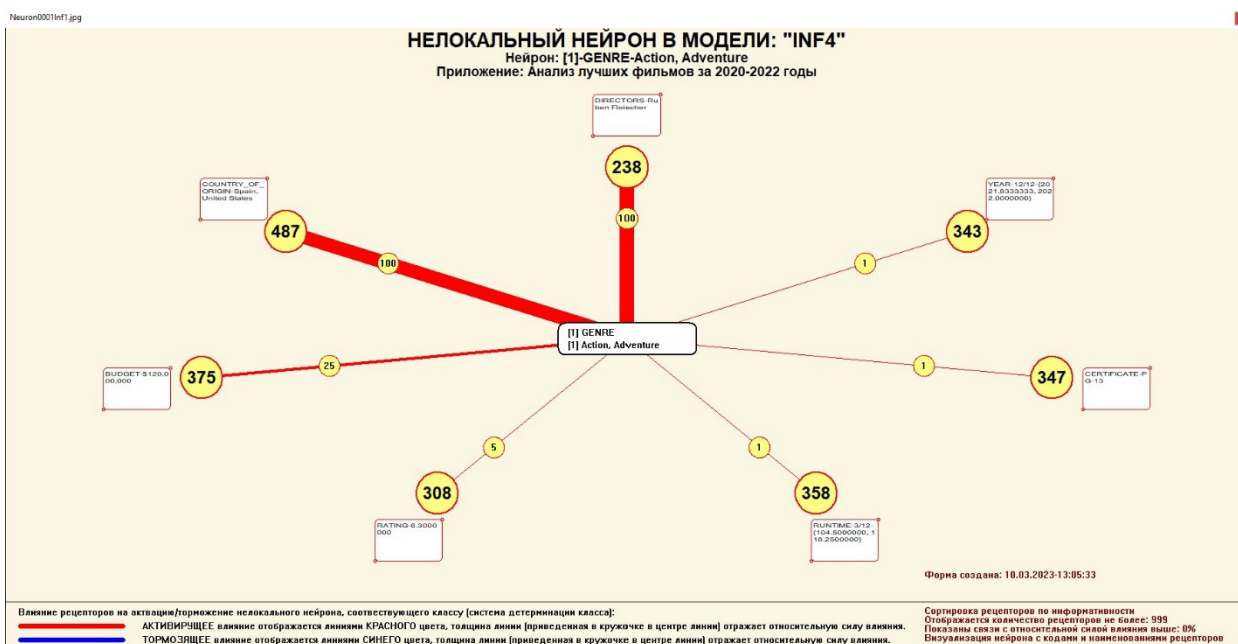


Рисунок 29 – Пример нелокального нейрона

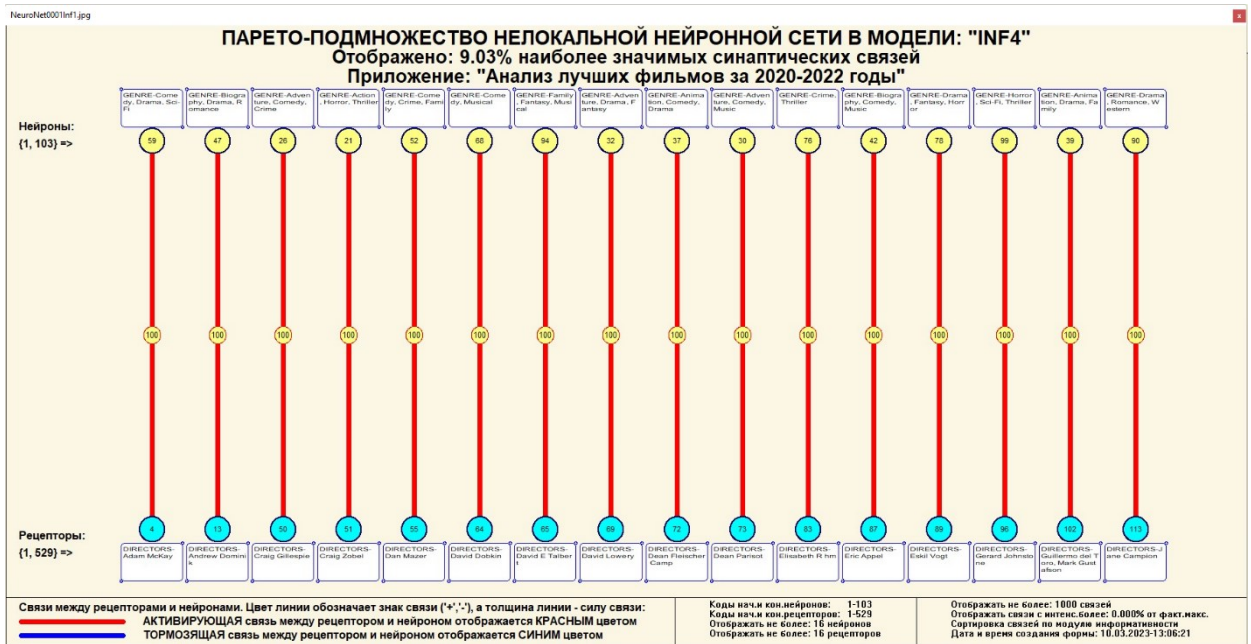


Рисунок 30 – Один слой нелокальной нейронной сети

2.3.6 3D-интегральные когнитивные карты

На рисунке 31 приведен фрагмент 3D-интегральной когнитивной карты, отражающий фрагмент СК-модели INF4.

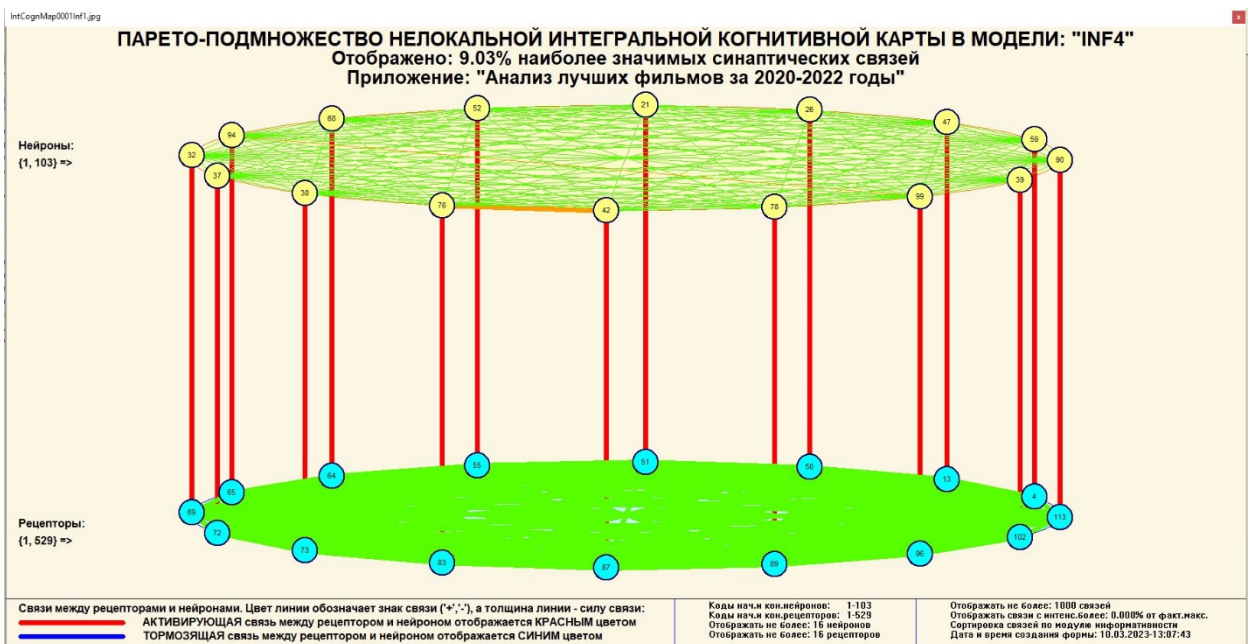


Рисунок 31 – 3D-интегральная когнитивная карта

ЗАКЛЮЧЕНИЕ

Цель работы достигнута, поставленная проблема решена, что показывает анализ результатов численного эксперимента предложенных и реализованных в системе «Aidos-X», решение поставленных задач является вполне эффективным.

При помощи системы «Aidos-X» результатом реализованной работы являются созданные статистические и системно-когнитивные модели, в которых на основе эмпирических данных были сформированы обобщенные образы классов, показывающие предполагаемый уровень заработной платы, и на основе этого, были достигнуты такие задачи, как идентификация, классификация и исследование моделируемой предметной области путем исследования ее модели.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Луценко Е.В. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос» / Е.В. Луценко, Е.К. Печурина, А.Э. Сергеев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2020. – №06(160). С. 95 – 114. – IDA [article ID]: 1602006009. – Режим доступа: <http://ej.kubagro.ru/2020/06/pdf/09.pdf>, 1,25 у.п.л.

2. Луценко Е.В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы «Эйдос-Х++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №07(101). С. 1367 – 1409. – IDA [article ID]: 1011407090. – Режим доступа: <http://ej.kubagro.ru/2014/07/pdf/90.pdf>, 2,688 у.п.л.

3. Луценко Е.В., Подсистема агломеративной когнитивной кластеризации классов системы «Эйдос» ("Эйдос-кластер"). Пат. № 2012610135 РФ. Заяв. № 2011617962 РФ 26.10.2011. Опубл. От 10.01.2012. – Режим доступа: <http://lc.kubagro.ru/aidos/2012610135.jpg>, 3,125 у.п.л.

4. Луценко Е.В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе

«Эйдос») / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №07(071). С. 528 – 576. – Шифр

Информрегистра: 0421100012\0253, IDA [article ID]: 0711107040. – Режим доступа: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>, 3,062 у.п.л.

5. Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК- анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №02(126). С. 1 – 32. – IDA [article ID]: 1261702001. – Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf>, 2 у.п.л.

6. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал

КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>, 1,562 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. Метод визуализации когнитивных функций – новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного

аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №03(067). С. 240 – 282. – Шифр Информрегистра: 0421100012\0077, IDA [article ID]: 0671103018. – Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.pdf>, 2,688 у.п.л.

9. Луценко Е.В. Сценарный АСК-анализ как метод разработки на основе эмпирических данных базисных функций и весовых коэффициентов для разложения в ряд функции состояния объекта или ситуации по теореме А.Н.Колмогорова (1957) / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2020. – №07(161). С. 76 – 120. – IDA [article ID]: 1612007009. – Режим доступа: <http://ej.kubagro.ru/2020/07/pdf/09.pdf>, 2,812 у.п.л.

10. Луценко Е.В. Детальный численный пример сценарного Автоматизированного системно-когнитивного анализа в интеллектуальной системе "Эйдос" / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2020. – №08(162). С. 273 – 355. – IDA [article ID]: 1622008020. – Режим доступа: <http://ej.kubagro.ru/2020/08/pdf/20.pdf>, 5,188 у.п.л.