

УДК 004.8

Автоматизированный системно-когнитивный анализ Топ-50 песен Spotify за 2019 год

Чувило Никита Владиславович
 студент факультета ПИ, группы ИТ2003
chuvilo-nikita@mail.ru

*Кубанский государственный аграрный университет
 имени И.Т.Трубилина, Краснодар, Россия*

Целью данной работы является изучение характеристик топ-50 песен Spotify за 2019 год. Достижение данной цели представляет большой личный интерес. Для нас это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-X++», а также получить зачет. Для достижения поставленной цели применяется Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос».

Ключевые слова: АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, АСК-АНАЛИЗ, СИСТЕМА «ЭЙДОС».

**Automated systemic cognitive analysis
 analysis of the top 50 songs on Spotify for 2019**

Chuvilo Nikita Vladislavovich
 student of the faculty of PI, group IT2003
chuvilo-nikita@mail.ru

Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia

The aim of this work is to study the characteristics of Spotify's top 50 songs for 2019. Achieving this goal is of great personal interest. For us, this will allow us to gain knowledge in the work with the universal cognitive analytical system "Eidos-HI" as well as get credit. To achieve this goal, Automated System and Cognitive Analysis (ASC-Analysis) and its software toolkit, the Eidos intelligent system, are used.

Key words: AUTOMATED SYSTEM-COGNITIVE ANALYSIS, ASC-ANALYSIS, "EIDOS" SYSTEM.

ОДЕРЖАНИЕ

ЗАДАЧА 1: КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	6
ЗАДАЧА 2: ПОДГОТОВКА ИСХОДНЫХ ДАННЫХ И ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ.....	7
ЗАДАЧА 3: СИНТЕЗ И ВЕРИФИКАЦИЯ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ И ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ ИЗ НИХ ДЛЯ РЕШЕНИЯ ЗАДАЧ	12
ЗАДАЧА 4: РЕШЕНИЕ РАЗЛИЧНЫХ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	17
Подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация)	17
Подзадача 4.2. Поддержка принятия решений (SWOT-анализ).....	19
Подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели	22
4.3.1. Когнитивные диаграммы классов.....	22
4.3.2. Агломеративная когнитивная кластеризация классов.....	24
4.3.3. Агломеративная когнитивная кластеризация значений факторов	25
4.3.4. Нелокальные нейроны и нелокальные нейронные сети.....	29
4.3.5. 3d-интегральные когнитивные карты.....	31
4.3.6. Когнитивные функции	32
7. ВЫВОДЫ	35
СПИСОК ЛИТЕРАТУРЫ	35

Введение

Целью данной работы является изучение характеристик топ-50 песен Spotify за 2019 год.

Достижение данной цели представляет большой личный интерес. Для нас это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-X++», а также получить зачет.

АСК-анализ предполагает, что для достижения поставленной цели необходимо решить следующие задачи, которые получаются путем декомпозиции цели и являются этапами ее достижения:

Задача 1: когнитивная структуризация предметной области.

Задача 2: подготовка исходных данных и формализация предметной области.

Задача 3: синтез и верификация статистических и системно-когнитивных моделей и выбор наиболее достоверной модели.

Задача 4: решение различных задач в наиболее достоверной модели:

- подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация);
- подзадача 4.2. Поддержка принятия решений;
- подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели (когнитивные диаграммы классов и значений факторов, агломеративная когнитивная кластеризация классов и значений факторов, нелокальные нейроны и нейронные сети, 3d-интегральные когнитивные карты, когнитивные функции).

Эти задачи, по сути, представляют собой **этапы** автоматизированного системно-когнитивного анализа (АСК-анализ), который поэтому и предлагается применить для их решения.

АСК-анализ представляет собой метод искусственного интеллекта, разработанный проф. Е.В. Луценко в 2002 году для решения широкого класса задач идентификации, прогнозирования, классификации, диагностики, поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели. АСК-анализ доведен до **инновационного** уровня благодаря тому, что имеет свой программный инструментарий – универсальную когнитивную аналитическую систему «Эйдос-X++» (система «Эйдос»).

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 (<http://lc.kubagro.ru/aidos/index.htm>), 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 (http://lc.kubagro.ru/aidos/_Aidos-X.htm), причем с актуальными исходными текстами (http://lc.kubagro.ru/_AIDOS-X.txt);
- 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» (есть акт внедрения системы «Эйдос» 1987 года) (<http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>);
- 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 (в настоящее время их 31 и 211, соответственно) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);
- поддерживает on-line среду накопления знаний и широко используется во всем мире (<http://aidos.byethost5.com/map5.php>);
- обеспечивает мультязычную 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 решения 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: http://lc.kubagro.ru/aidos/aidos18_LLS/aidos18_LLS.pdf);

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

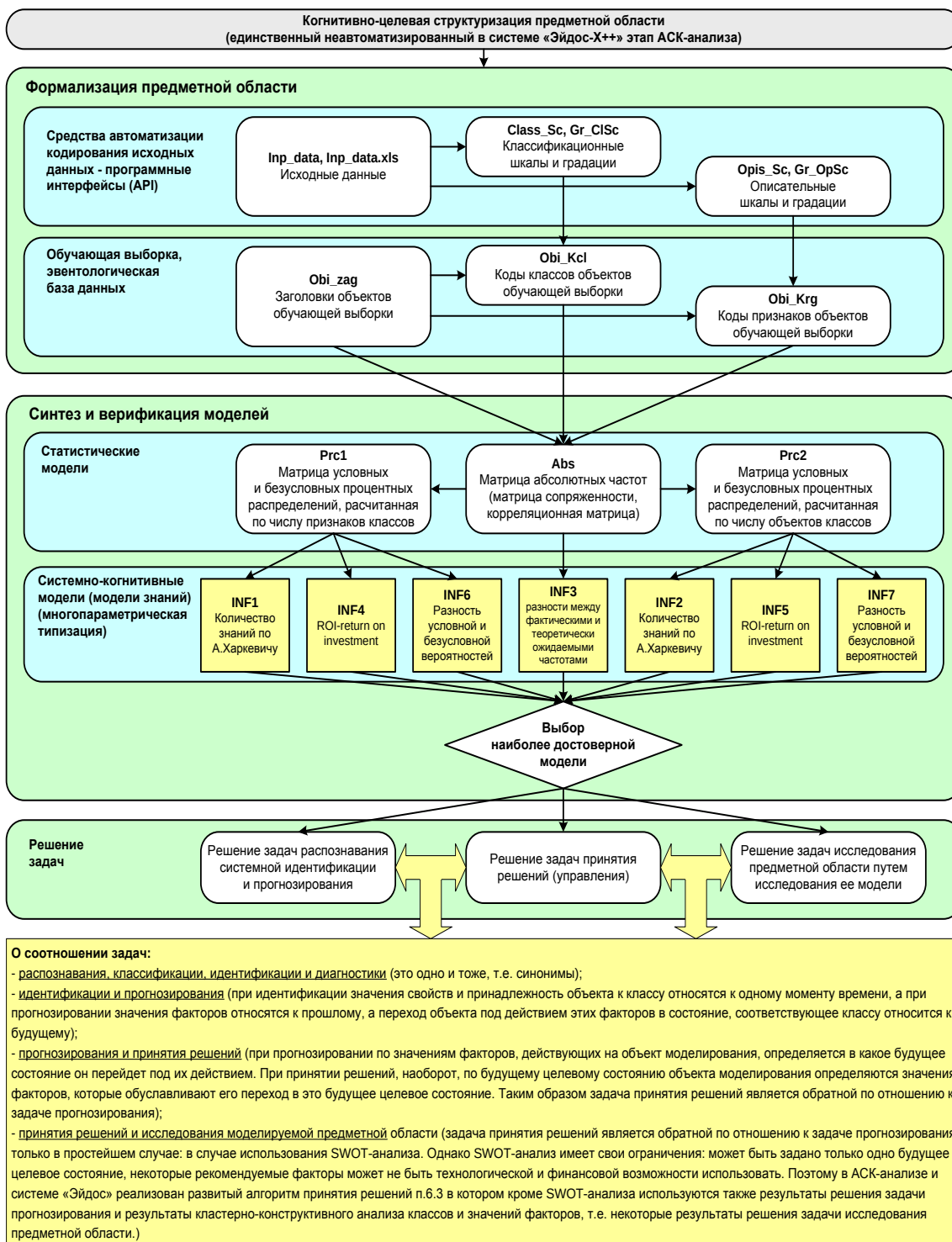
- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторностей всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ) 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.

В чем сила подхода, реализованного в системе Эйдос? В том, что она реализует подход, эффективность которого не зависит от того, что мы думаем о предметной области и думаем ли вообще. Она формирует модели непосредственно на основе эмпирических данных, а не на основе наших представлений о механизмах реализации закономерностей в этих данных. Именно поэтому Эйдос-модели эффективны даже если наши представления о предметной области ошибочны или вообще отсутствуют.

В этом и слабость этого подхода, реализованного в системе Эйдос. Модели системы Эйдос - это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они не отражают причинно-следственного механизма детерминации, а только сам факт и характер детерминации. Содержательное объяснение этих эмпирических закономерностей формулируется уже на теоретическом уровне познания в теоретических научных законах.

Всем этим и обусловлен выбор АСК-анализа и его программного инструментария – интеллектуальной системы «Эйдос» в качестве метода и инструмента решения поставленной проблемы и достижения цели работы (рисунок 1).

**Последовательность обработки данных, информации и знаний в системе «Эйдос»,
повышение уровня системности данных, информации и знаний,
повышение уровня системности моделей**



**Рисунок 1. Последовательность решения задач
в АСК-анализе и системе «Эйдос»**

Рассмотрим решение поставленных задач в подробном численном примере.

Задача 1: когнитивная структуризация предметной области

На этапе когнитивно-целевой структуризации предметной области мы неформализуемым путем решаем на качественном уровне, что будем рассматривать в качестве факторов, действующих на моделируемый объект (причин), а что в качестве результатов действия этих факторов (последствий).

При этом необходимо отметить, что системно-когнитивные модели (СК-модели) отражают лишь сам факт наличия зависимостей между значениями факторов и результатами их действия. Но они не отражают причин и механизмов такого влияния.

Это значит:

– во-первых, что содержательная интерпретация СК-моделей – это компетенция специалистов-экспертов хорошо разбирающихся в данной предметной области. Иногда встречается ситуация, когда и то, что на первый взгляд является причинами, и то, что казалось бы является их последствиями, на самом деле является последствиями неких глубинных причин, которых мы не видим и никоим образом непосредственно не отражаем в модели;

– во-вторых, даже если содержательной интерпретации не разработано, то в принципе это не исключает возможности пользоваться ими на практике для достижения заданных результатов и поставленных целей, т.е. для управления.

В данной работе в качестве классификационных шкал выберем популярность (таблица 1), а в качестве факторов, влияющих на популярность – ее жанр, BPM, энергия, танцевальность, громкость, живучесть, длительность, акустика, речевость.(таблица 2):

Таблица 1 – Классификационная шкала

Код	Наименование
1	Popularity

Таблица 2 – Описательные шкалы

Код	Наименование
1	genre
2	BPM
3	energy
4	danceability
5	loudness
6	liveness
7	length
8	acousticness
9	speechiness

Задача 2: подготовка исходных данных и формализация предметной области

Исходные данные для данной статьи (рисунок 2) получены из Spotify.

Ob	Track.Name	Artist.Name	Genre	Beats.Per.Minute	Energy	Danceability	Loudness..dB..	Liveness	Valence.	Length.	Acousticness..	Speechiness.	Popularity
1	Secorita	Shawn Mendes	canadian pop	117	55	76	-6	8	75	191	4	3	79
2	China	Anuel AA	reggaeton flow	105	81	79	-4	8	61	302	8	9	92
3	boyfriend (with Social House)	Ariana Grande	dance pop	190	80	40	-4	16	70	186	12	46	85
4	Beautiful People (feat. Khalid)	Ed Sheeran	pop	93	65	64	-8	8	55	198	12	19	86
5	Goodbyes (Feat. Young Thug)	Post Malone	dfw rap	150	65	58	-4	11	18	175	45	7	94
6	I Don't Care (with Justin Bieber)	Ed Sheeran	pop	102	68	80	-5	9	84	220	9	4	84
7	Ransom	Lil Tecca	trap music	180	64	75	-6	7	23	131	2	29	92
8	How Do You Sleep?	Sam Smith	pop	111	68	48	-5	8	35	202	15	9	90
9	Old Town Road - Remix	Lil Nas X	country rap	136	62	88	-6	11	64	157	5	10	87
10	bad guy	Billie Eilish	electropop	135	43	70	-11	10	56	194	33	38	95
11	Callaita	Bad Bunny	reggaeton	176	62	61	-5	24	24	251	60	31	93
12	Loco Contigo (feat. J. Balvin & Tyga)	DJ Snake	dance pop	96	71	82	-4	15	38	185	28	7	86
13	Someone You Loved	Lewis Capaldi	pop	110	41	50	-6	11	45	182	75	3	88
14	Otro Trago - Remix	Sech	panamanian pop	176	79	73	-2	6	76	288	7	20	87
15	Money In The Grave (Drake ft. Rick Ross)	Drake	canadian hip hop	101	50	83	-4	12	10	205	10	5	92
16	No Guidance (feat. Drake)	Chris Brown	dance pop	93	45	70	-7	16	14	261	12	15	82
17	LA CANCIYN	J Balvin	latin	176	65	75	-6	11	43	243	15	32	90
18	Sunflower - Spider-Man: Into the Spider-Verse	Post Malone	dfw rap	90	48	76	-6	7	91	158	56	5	91
19	Lalala	Y2K	canadian hip hop	130	39	84	-8	14	50	161	18	8	88
20	Truth Hurts	Lizzo	escape room	158	62	72	-3	12	41	173	11	11	91
21	Piece Of Your Heart	MEDUZA	pop house	124	74	68	-7	7	63	153	4	3	91
22	Panini	Lil Nas X	country rap	154	59	70	-6	12	48	115	34	8	91
23	No Me Conoce - Remix	Jhay Cortez	reggaeton flow	92	79	81	-4	9	58	309	14	7	83
24	Soltera - Remix	Lunay	latin	92	78	80	-4	44	80	266	36	4	91
25	bad guy (with Justin Bieber)	Billie Eilish	electropop	135	45	67	-11	12	68	195	25	30	89
26	If I Can't Have You	Shawn Mendes	canadian pop	124	82	69	-4	13	87	191	49	6	70
27	Dance Monkey	Tones and I	australian pop	98	59	82	-6	18	54	210	69	10	83
28	It's You	Ali Gatie	canadian hip hop	96	46	73	-7	19	40	213	37	3	89
29	Con Calma	Daddy Yankee	latin	94	86	74	-3	6	66	193	11	6	91
30	QUE PRETENDES	J Balvin	latin	93	79	64	-4	36	94	222	3	25	89
31	Takeaway	The Chainsmokers	edm	85	51	29	-8	10	36	210	12	4	84
32	7 rings	Ariana Grande	dance pop	140	32	78	-11	9	33	179	59	33	89
33	don't call me up	Maluma	reggaeton	96	71	78	-5	9	68	176	22	28	89
34	The London (feat. J. Cole & Travis Scott)	Young Thug	atl hip hop	98	59	80	-7	13	18	200	2	15	89
35	Never Really Over	Katy Perry	dance pop	100	88	77	-5	32	39	224	19	6	89
36	Summer Days (feat. Macklemore & Patrick Stump of Fall Out Boy)	Martin Garrix	big room	114	72	66	-7	14	32	164	18	6	89
37	Otro Trago	Sech	panamanian pop	176	70	75	-5	11	62	226	14	34	91
38	Antisocial (with Travis Scott)	Ed Sheeran	pop	152	82	72	-5	36	91	162	13	5	87
39	Sucker	Jonas Brothers	boy band	138	73	84	-5	11	95	181	4	6	80
40	fuck. i'm lonely (with Anne-Marie) - from "13 Reasons Why: Season 3"	Lauv	dance pop	95	56	81	-6	6	68	199	48	7	78
41	Higher Love	Kygo	edm	104	68	69	-7	10	40	228	2	3	88
42	You Need To Calm Down	Taylor Swift	dance pop	85	68	77	-6	7	73	171	1	5	90
43	Shallow	Lady Gaga	dance pop	96	39	57	-6	23	32	216	37	3	87
44	Talk	Khalid	pop	136	40	90	-9	6	35	198	5	13	84
45	Con Altura	ROSALHA	r&b en espanol	98	69	88	-4	5	75	162	39	12	88
46	One Thing Right	Marshmello	brostep	88	62	66	-2	58	44	182	7	5	88
47	Te Robarí	Nicky Jam	latin	176	75	67	-4	8	80	202	24	6	88
48	Happier	Marshmello	brostep	100	79	69	-3	17	67	214	19	5	88
49	Call You Mine	The Chainsmokers	edm	104	70	59	-6	41	50	218	23	3	88
50	Cross Me (feat. Chance the Rapper & PnB Rock)	Ed Sheeran	pop	95	79	75	-6	7	61	206	21	12	82

Рисунок 2 – Исходные данные для ввода в систему «Эйдос»

Затем с параметрами, показанными на рисунке 3, запустим режим 2.3.2.2 системы «Эйдос», представляющий собой автоматизированный программный интерфейс (API) с внешними данными табличного типа.

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 диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал: 14
Конечный столбец классификационных шкал: 14

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал: 2
Конечный столбец описательных шкал: 13

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 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":

Интерпретация TXT-полей классов:

Значения полей текстовых классификационных шкал файла исходных данных "Inp_data" рассматриваются как целое

Интерпретация TXT-полей признаков:

Значения полей текстовых описательных шкал файла исходных данных "Inp_data" рассматриваются как целое

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

2.3.2.2. Задание размерности модели системы "ЭЙДОС-X++"

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ

Суммарное количество градаций классификационных и описательных шкал: [7 x 136]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	1	7	7,00	9	27	3,00
Текстовые	0	0	0,00	3	109	36,33
ВСЕГО:	1	7	7,00	12	136	11,33

Задайте число интервалов (градаций) в шкале:

В классификационных шкалах: 7 В описательных шкалах: 3

Пересчитать шкалы и градации Параметры числ. шкал и градаций Выйти на создание модели

2.3.2.2. Параметры классификационных и описательных шкал и градаций

ПАРАМЕТРЫ ШКАЛ И ГРАДАЦИЙ С АДАПТИВНЫМИ ГРАНИЦАМИ И ПРИМЕРНО РАВНЫМ КОЛИЧЕСТВОМ НАБЛЮДЕНИЙ ПО ГРАДАЦИЯМ с коррекцией ошибки округления числа наблюдений по интервалу градации при переходе к следующей градации

КЛАССИФИКАЦИОННАЯ ШКАЛА: код: [1], наим.: "POPULARITY", набл. на шкалу (всего): 50, тип/число градаций в шкале: "Равное число событий в интервалах"

1 Наим. градации: 1/7-[70.0000000, 83.0000000], размер интервала=13.0000000, расч./факт. число наблюдений на градацию: 7/7
 2 Наим. градации: 2/7-[83.0000000, 86.0000000], размер интервала= 3.0000000, расч./факт. число наблюдений на градацию: 7/7
 3 Наим. градации: 3/7-[86.0000000, 88.0000000], размер интервала= 2.0000000, расч./факт. число наблюдений на градацию: 7/7
 4 Наим. градации: 4/7-[88.0000000, 89.0000000], размер интервала= 1.0000000, расч./факт. число наблюдений на градацию: 7/7
 5 Наим. градации: 5/7-[89.0000000, 90.0000000], размер интервала= 1.0000000, расч./факт. число наблюдений на градацию: 7/7
 6 Наим. градации: 6/7-[90.0000000, 91.0000000], размер интервала= 1.0000000, расч./факт. число наблюдений на градацию: 7/7
 7 Наим. градации: 7/7-[91.0000000, 95.0000000], размер интервала= 4.0000000, расч./факт. число наблюдений на градацию: 6/8

ОПИСАТЕЛЬНАЯ ШКАЛА: код: [1], наим.: "TRACK.NAME", тип/число градаций в шкале: "Равное число событий в интервалах"/50

1 Наим. градации: 1/50-7 rings
 2 Наим. градации: 2/50-Antisocial (with Travis Scott)
 3 Наим. градации: 3/50-bad guy
 4 Наим. градации: 4/50-bad guy (with Justin Bieber)
 5 Наим. градации: 5/50-Beautiful People (feat. Khalid)
 6 Наим. градации: 6/50-boyfriend (with Social House)
 7 Наим. градации: 7/50-Call You Mine

Ok

Рисунок 3. Экранные форма программного интерфейса (API) 2.3.2.2 системы «Эйдос» с внешними данными табличного типа

Обратим внимание на то, что заданы адаптивные интервалы, учитывающее неравномерность распределения данных по диапазону значений, что важно при относительно небольшом числе наблюдений. Если бы интервалы были заданы равными по величине, то в них бы учитывалось сильно отличающееся число наблюдений, а в некоторых интервалах их бы могло не оказаться вовсе. В описательных шкалах задано 3 числовых интервальных значения.

На рисунке 4 приведен Help данного режима, в котором объясняется принцип организации таблицы исходных данных для данного режима. Здесь же обратим внимание на то, что в таблице 3 значения параметров могут быть представлены как числовыми, так и текстовыми значениями.

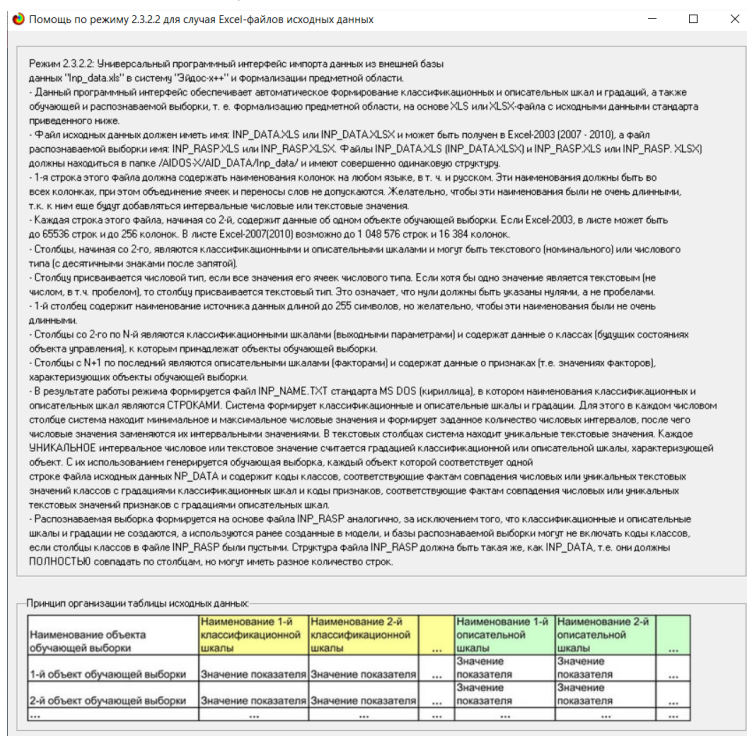


Рисунок 4. Экранная форма HELP программного интерфейса (API) 2.3.2.2

В результате работы режима сформирована 1 классификационная шкала с суммарным количеством градаций (классов) 7 (таблица 3) и 15 описательных шкал с суммарным числом градаций 136 (таблица 4). С использованием классификационных и описательных шкал и градаций исходные данные (рисунок 2) были закодированы и в результате получена обучающая выборка (таблица 5):

Таблица 3 – Классификационные шкалы и градации (класс оружия)

Код	Название
1	1/7-{70.0000000, 83.0000000}
2	2/7-{83.0000000, 86.0000000}
3	3/7-{86.0000000, 88.0000000}
4	4/7-{88.0000000, 89.0000000}
5	5/7-{89.0000000, 90.0000000}
6	6/7-{90.0000000, 91.0000000}
7	7/7-{91.0000000, 95.0000000}

Таблица 4 – Описательные шкалы и градации
(характеристики оружия)

Код	Название
1.	1/50-7 rings
2.	2/50-Antisocial(withTravisScott)
3.	3/50-badguy
4.	4/50-badguy(withJustinBieber)
5.	5/50-BeautifulPeople(feat.Khalid)
6.	6/50-boyfriend(withSocialHouse)
7.	7/50-CallYouMine
8.	8/50-Callaita
9.	9/50-China
10.	10/50-ConAltura
11.	11/50-ConCalma
12.	12/50-CrossMe(feat.ChancetheRapper&PnBRock)
13.	13/50-DanceMonkey
14.	14/50-don'tcallmeup
15.	15/50-fuck.i'mlonely(withAnne-Marie)-from"13ReasonsWhySeason3"
16.	16/50-Goodbyes(Feat.YoungThug)
17.	17/50-Happier
18.	18/50-HigherLove
19.	19/50-HowDoYouSleep?
20.	20/50-IDon'tCare(withJustinBieber)
21.	21/50-IfICan'tHaveYou
22.	22/50-It'sYou
23.	23/50-LACANCI*N
24.	24/50-Lalala
25.	25/50-LoCoContigo(feat.J.Balvin&Tyga)
26.	26/50-MoneyInTheGrave(Drakeft.RickRoss)
27.	27/50-NeverReallyOver
28.	28/50-NoGuidance(feat.Drake)
29.	29/50-NoMeConoce-Remix
30.	30/50-OldTownRoad-Remix
31.	31/50-OneThingRight
32.	32/50-OtroTrago-Remix
33.	33/50-OtroTrago
34.	34/50-Panini
35.	35/50-PieceOfYourHeart
36.	36/50-QUEPRETENDES
37.	37/50-Ransom
38.	38/50-Se6orita
39.	39/50-Shallow
40.	40/50-Soltera-Remix
41.	41/50-SomeoneYouLoved
42.	42/50-Sucker
43.	43/50-SummerDays
44.	44/50-Sunflower-Spider-ManIntotheSpider-Verse
45.	45/50-Takeaway
46.	46/50-Talk
47.	47/50-TeRoBar©
48.	48/50-TheLondon(feat.J.Cole&TravisScott)
49.	49/50-TruthHurts
50.	50/50-YouNeedToCalmDown
51.	1/38-AliGatie
52.	2/38-AnuelAA
53.	3/38-ArianaGrande
54.	4/38-BadBunny
55.	5/38-BillieEilish
56.	6/38-ChrisBrown
57.	7/38-DaddyYankee
58.	8/38-DJSnake
59.	9/38-Drake
60.	10/38-EdSheeran
61.	11/38-JBalvin
62.	12/38-JhayCortez
63.	13/38-JonasBrothers
64.	14/38-KatyPerry
65.	15/38-Khalid
66.	16/38-Kygo
67.	17/38-LadyGaga

68.	18/38-Lauv
69.	19/38-LewisCapaldi
70.	20/38-LilNasX
71.	21/38-LilTecca
72.	22/38-Lizzo
73.	23/38-Lunay
74.	24/38-Maluma
75.	25/38-Marshmello
76.	26/38-MartinGarrix
77.	27/38-MEDUZA
78.	28/38-NickyJam
79.	29/38-PostMalone
80.	30/38-ROSALÍA
81.	31/38-SamSmith
82.	32/38-Sech
83.	33/38-ShawnMendes
84.	34/38-TaylorSwift
85.	35/38-TheChainsmokers
86.	36/38-TonesandI
87.	37/38-Y2K
88.	38/38-YoungThug
89.	1/21-atlhiphop
90.	2/21-australianpop
91.	3/21-bigroom
92.	4/21-boyband
93.	5/21-brostep
94.	6/21-canadianhiphop
95.	7/21-canadianpop
96.	8/21-countryrap
97.	9/21-dancepop
98.	10/21-dfwrap
99.	11/21-edm
100.	12/21-electropop
101.	13/21-escaperoom
102.	14/21-latin
103.	15/21-panamanianpop
104.	16/21-pop
105.	17/21-pophouse
106.	18/21-r&b&spanol
107.	19/21-reggaeton
108.	20/21-reggaetonflow
109.	21/21-trapmusic
110.	1/3-{85.000000096.0000000}
111.	2/3-{96.0000000130.0000000}
112.	3/3-{130.0000000190.0000000}
113.	1/3-{32.000000059.0000000}
114.	2/3-{59.000000071.0000000}
115.	3/3-{71.000000088.0000000}
116.	1/3-{29.000000069.0000000}
117.	2/3-{69.000000077.0000000}
118.	3/3-{77.000000090.0000000}
119.	1/3-{-11.0000000-6.0000000}
120.	2/3-{-6.0000000-5.0000000}
121.	3/3-{-5.0000000-2.0000000}
122.	1/3-{5.00000009.0000000}
123.	2/3-{9.000000013.0000000}
124.	3/3-{13.000000058.0000000}
125.	1/3-{10.000000040.0000000}
126.	2/3-{40.000000066.0000000}
127.	3/3-{66.000000095.0000000}
128.	1/3-{115.0000000182.0000000}
129.	2/3-{182.0000000210.0000000}
130.	3/3-{210.0000000309.0000000}
131.	1/3-{1.000000011.0000000}
132.	2/3-{11.000000023.0000000}
133.	3/3-{23.000000075.0000000}
134.	1/3-{3.00000005.0000000}
135.	2/3-{5.000000011.0000000}
136.	3/3-{11.000000046.0000000}

2.3.1. Ручной ввод-корректировка обучающей выборки. Текущая модель: "INF1"

Код объекта	Наименование объекта	Дата	Время
1	1		
2	2		
3	3		
4	4		
5	5		
6	6		
7	7		
8	8		
9	9		
10	10		

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4
1	1	0	0	0
2				
3				
4				
5				
6				
7				
8				
9				
10				

Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	38	83	95	111	113	117	119
1	122	127	129	131	134	0	0

Помощь Скопировать обуч.выб.в расп. Добавить объект Добавить классы Добавить признаки Удалить объект Удалить классы Удалить признаки Очистить БД

Рисунок 4а. Обучающая выборка (фрагмент)

Задача 3: синтез и верификация статистических и системно-когнитивных моделей и выбор наиболее достоверной из них для решения задач

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) моделей осуществляется в режиме 3.5 системы «Эйдос» (рисунок 5).

3.5. Выбор моделей для синтеза и верификации

Задайте модели для синтеза и верификации:

Статистические базы:

- ☒ 1. ABS - частный критерий: количество встреч сочетаний: "класс-признак" у объектов обуч.выборки
- ☒ 2. PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов i-го класса
- ☒ 3. PRC2 - частный критерий: условная вероятность i-го признака у объектов i-го класса

Системно-когнитивные модели (базы знаний):

- ☒ 4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1
- ☒ 5. INF2 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2
- ☒ 6. INF3 - частный критерий: χ^2 -квдрат, разности между фактическими и ожидаемыми абс. частотами
- ☒ 7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1
- ☒ 8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2
- ☒ 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1
- ☒ 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2

Текущая модель:

- ☐ 1. ABS
- ☐ 2. PRC1
- ☐ 3. PRC2
- ☒ 4. INF1
- ☐ 5. INF2
- ☐ 6. INF3
- ☐ 7. INF4
- ☐ 8. INF5
- ☐ 9. INF6
- ☐ 10. INF7

Параметры копирования обучающей выборки в распознаваемую:

Какие объекты обуч.выборки копировать:

- ☒ Копировать всю обучающую выборку
- ☐ Копировать только текущий объект
- ☐ Копировать каждый N-й объект
- ☐ Копировать N случайных объектов
- ☐ Копировать все объекты от N1 до N2
- ☐ Вообще не менять распознаваемую выборку

Удалять из обуч.выборки скопированные объекты:

- ☒ Не удалять
- ☐ Удалять

Пояснение по алгоритму верификации

Измеряется внутренняя достоверн. модели

Для каждой заданной модели выполнять:

- ☒ Синтез и верификацию
- ☐ Только верификацию
- ☐ Только синтез

На каком процессоре выполнять расчеты:

- ☐ CPU
- ☒ GPU

Уменьшение размеров базы данных результатов распознавания: Rasp.dbf

Расчетный размер БД результатов распознавания Rasp.dbf равен 7000 байт, т.е.: 0.0003260 % от MAX-возможного, (от 2Гб)

Задайте, сколько % от исходной БД Rasp.dbf оставить, удаляя наименее достоверные результаты распознавания:

Ok Cancel

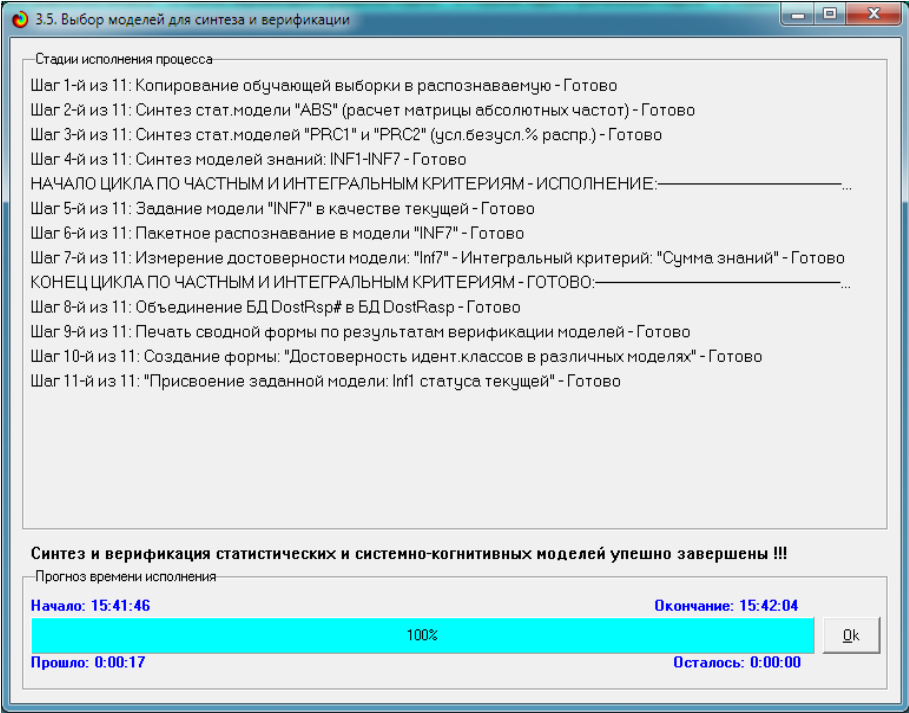


Рисунок 5. Экранные формы режима синтеза и верификации статистических и системно-когнитивных моделей системы «Эйдос»

Обратим внимание на то, что на рисунке 5 в правом нижнем углу окна задана опция: «Расчеты проводить на графическом процессоре (GPU)».

Из рисунка 5 видно, что весь процесс синтеза и верификации моделей занял 17 секунд. Отметим, что при синтезе и верификации моделей использовался графический процессор (GPU) видеокарты. На центральном процессоре (CPU) выполнение этих операций занимает значительно большее время (на некоторых задачах это происходит в десятки, сотни и даже тысячи раз дольше). Таким образом, неграфические вычисления на графических процессорах видеокарты делает возможной обработку больших объемов исходных данных за разумное время. В процесс синтеза и верификации моделей осуществляется также расчет 10 выходных форм, на что уходит более 99% времени исполнения.

Фрагменты самих созданных статистических и системно-когнитивных моделей (СК-модели) приведены на рисунках 6, 7:

3.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абсолютными"										
Код признака	Наименование описательных параметров и градаций	1. POPULARITY 1/7 (70.000000, 73.571429)	2. POPULARITY 2/7 (73.571429, 77.142857)	3. POPULARITY 3/7 (77.142857, 80.7142857)	4. POPULARITY 4/7 (80.7142857, 84.2957143)	5. POPULARITY 5/7 (84.2957143, 87.8571429)	6. POPULARITY 6/7 (87.8571429, 91.4295714)	7. POPULARITY 7/7 (91.4295714, 95.000000)	Средне	
119	DANCEABILITY-3/3 (89.6666667, 90.0000000)	-0.640		1.080	1.520	-0.480	-1.640	0.160	0.000	
120	LOUDNESS_DB-1/3 (11.000000, 0.0000000)	-0.140		-0.420	1.020	0.020	-0.640	0.160	0.000	
121	LOUDNESS_DB-2/3 (0.000000, 5.0000000)	-0.540		1.380	0.220	-0.780	0.960	-1.240	0.000	
122	LIVENESS_DB-3/3 (5.000000, 22.6666667)	0.680		-0.960	-1.240	0.760	-0.320	1.080	0.000	
123	LIVENESS-1/3 (5.000000, 22.6666667)	0.160		0.480	1.120	-0.880	-0.840	-0.040	0.000	
124	LIVENESS-2/3 (22.6666667, 40.3333333)	-0.100		-0.300	-0.700	1.300	-0.600	0.400	0.000	
125	LIVENESS-3/3 (40.333333, 58.0000000)	-0.060		-0.180	-0.420	-0.420	1.440	-0.360	0.000	
126	VALENCE-1/3 (10.000000, 38.3333333)	-0.260		-0.780	1.180	0.180	-2.760	2.440	0.000	
127	VALENCE-2/3 (38.333333, 66.6666667)	-0.400		-1.200	0.200	-0.800	2.600	-0.400	0.000	
128	VALENCE-3/3 (66.666667, 95.0000000)	0.660		1.980	-1.380	0.620	0.160	-2.040	0.000	
129	LENGTH-1/3 (115.000000, 175.6666667)	-0.280		-0.840	-1.960	0.040	2.720	0.320	0.000	
130	LENGTH-2/3 (175.666667, 244.3333333)	0.400		1.200	0.800	-0.200	-0.600	-1.600	0.000	
131	LENGTH-3/3 (244.333333, 309.0000000)	-0.120		-0.360	1.160	0.160	-2.120	1.280	0.000	
132	ACQUSSIONESS-1/3 (1.000000, 25.6666667)	-0.700		-0.100	1.100	0.100	0.800	-1.200	0.000	
133	ACQUSSIONESS-2/3 (25.666667, 50.3333333)	0.800		0.400	-1.400	0.600	-1.200	0.800	0.000	
134	ACQUSSIONESS-3/3 (50.333333, 75.0000000)	-0.100		-0.300	0.300	-0.700	0.400	0.400	0.000	
135	SPEECHNESS-1/3 (1.000000, 17.3333333)	0.240		0.720	1.680	-1.320	0.240	-1.560	0.000	
136	SPEECHNESS-2/3 (17.333333, 31.6666667)	-0.140		-0.420	-0.980	1.020	-0.640	1.160	0.000	
137	SPEECHNESS-3/3 (31.666667, 46.0000000)	-0.100		-0.300	-0.700	0.300	0.400	0.400	0.000	
Средне		0.000		0.000	0.000	0.000	0.000	0.000	0.000	
Средне-квадратичное отклонение		0.248		0.424	0.632	0.494	0.810	0.592	0.000	

Рисунок 6. Матрица абсолютных частот (фрагмент)

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

Год	Наименование описательной шкалы и градаций	1 POPULARITY 1/7 (70.0, 83.0)	2 POPULARITY 2/7 (83.0, 88.0)	3 POPULARITY 3/7 (88.0, 89.0)	4 POPULARITY 4/7 (89.0, 89.0)	5 POPULARITY 5/7 (89.0, 90.0)	6 POPULARITY 6/7 (90.0, 91.0)	7 POPULARITY 7/7 (91.0, 95.0)	Средн	Среднее
119	LOUDNESS: 06 -1/3(11.000000, 4.000000)	0.840	-0.120	-0.240	0.840	0.440	-0.440	-1.120		
120	LOUDNESS: 08 -2/3(4.000000, 6.000000)	-0.280	0.040	-0.920	0.720	0.920	-0.120	0.040		
121	LOUDNESS: 08 -3/3(5.000000, 3.000000)	-0.560	0.080	1.160	-1.560	-0.960	0.760	1.000	0.000	0.000
122	LIVENESS: 1/3(5.000000, 3.000000)	0.960	0.720	-1.960	-1.040	0.840	0.340	-0.280	0.000	0.000
123	LIVENESS: 2/3(9.000000, 13.000000)	-0.400	-0.800	-0.600	-0.400	0.100	0.900	1.200		
124	LIVENESS: 3/3(13.000000, 19.000000)	-0.540	0.090	2.140	1.440	-0.940	-1.240	-0.920	0.000	0.000
125	VALENCE: 1/3(10.000000, 40.000000)	-1.560	1.080	-1.840	2.440	0.040	-2.240	2.080		
126	VALENCE: 2/3(40.000000, 66.000000)	0.280	-1.040	0.920	-2.720	-0.020	2.620	-0.040		
127	VALENCE: 3/3(66.000000, 95.000000)	1.280	-0.040	0.920	0.280	-0.020	-0.380	-2.040		
128	LENGTH: 1/3(115.000000, 182.000000)	-1.720	-2.040	1.920	0.280	-0.020	1.620	-0.040		
129	LENGTH: 2/3(182.000000, 210.000000)	2.280	2.960	-3.080	-0.720	-0.020	-1.380	-0.040	0.000	0.000
130	LENGTH: 3/3(210.000000, 309.000000)	-0.560	-0.920	1.160	0.440	0.040	-0.240	0.080		
131	ACOUSTICNESS: 1/3(11.000000, 11.000000)	-0.720	-0.040	-0.080	-0.720	-0.020	0.620	0.960		
132	ACOUSTICNESS: 2/3(11.000000, 23.000000)	0.440	1.080	0.160	0.440	1.040	-1.240	-1.920		
133	ACOUSTICNESS: 3/3(23.000000, 75.000000)	0.280	-1.040	-0.080	0.280	-1.020	0.620	0.960		
134	SPEECHNESS: 1/3(3.000000, 5.000000)	-1.560	0.080	3.160	-1.560	0.040	0.760	-0.920	0.000	0.000
135	SPEECHNESS: 2/3(5.000000, 11.000000)	2.280	-1.040	-1.080	-0.720	-0.020	0.620	-0.040		
136	SPEECHNESS: 3/3(11.000000, 46.000000)	-0.720	0.960	-2.080	2.280	-0.020	-1.380	0.960		
Средн		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Среднее		0.709	0.605	0.822	0.638	0.385	0.647	0.586		
Среднеквадратичное отклонение										

Рисунок 7. Модель INF3 (фрагмент)

Отметим, что в АСК-анализе и СК-моделях степень выраженности различных свойств объектов наблюдения рассматривается с единственной точки зрения: с точки зрения того, какое *количество информации* содержится в них о том, к каким обобщающим категориям (классам) будут принадлежать или не принадлежать эти объекты. Поэтому не играет никакой роли, в каких единицах измерения измеряются те или иные свойства объектов наблюдения, а также в каких единицах измеряются результаты влияния этих свойств, натуральных, в процентах или стоимостных. Это и есть решение проблемы сопоставимости в АСК-анализе и системе «Эйдос», отличающее их от других интеллектуальных технологий.

Верификация статистических и системно-когнитивных моделей

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных положительных и отрицательных, а также ложных положительных и отрицательных решений по F-мере Ван Ризбергера, а также по критериям L1-L2-мерам проф. Е.В.Луценко, которые предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры. В режиме 3.4 системы «Эйдос» изучается достоверность каждой частной модели в соответствии с этими мерами достоверности (рисунок 8).

3.4. Обобщенная форма по достоверности при разнотип. Текущая модель: "INF3"

Наименование модели и частного критерия	Интегральный критерий	3 макс. макс. чл (FP)	Число макс. отрицательных решений (FN)	Точность модели	Полная модель	F-мера Ван Ризбергера	Средн. макс. долей с. истинно-поло. решений (ST)	Средн. макс. долей с. истинно-отриц. решений (ST)	Средн. макс. долей с. ложно-поло. решений (FP)	Средн. макс. долей с. ложно-отриц. решений (FN)	5-Точность модели	5-Полная модель	L1-мера проф. Е.В.Луценко
1. ABS - частный критерий: количество встреч соединений, "хлос"	Корреляция абс. частот с обр...	300		0.143	1.000	0.250	38.570		138.373	0.218	1.000	0.339	0.339
1. ABS - частный критерий: количество встреч соединений, "хлос"	Средн. абс. частот по прина...	300		0.143	1.000	0.250	38.570		138.373	0.218	1.000	0.339	0.339
2. PRIC1 - частный критерий: усл. вероятность его признака сред.	Корреляция усл. частот с в. пр...	300		0.143	1.000	0.250	37.800		140.212	0.191	1.000	0.301	0.301
3. PRIC2 - частный критерий: усл. вероятность его признака	Корреляция усл. частот с в. пр...	300		0.143	1.000	0.250	38.570		138.373	0.218	1.000	0.339	0.339
3. PRIC2 - частный критерий: усл. вероятность его признака	Средн. усл. частот по при...	300		0.143	1.000	0.250	37.800		140.212	0.191	1.000	0.301	0.301
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	21	4	0.687	0.920	0.756	19.897	106.640	2.321	0.465	0.896	0.977	0.935
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	136		0.269	1.000	0.404	30.611	24.296	18.508	0.623	1.000	0.769	0.769
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	21	4	0.687	0.920	0.756	19.897	106.640	2.321	0.465	0.896	0.977	0.935
6. INF3 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	136		0.269	1.000	0.404	30.611	24.296	18.508	0.623	1.000	0.769	0.769
6. INF3 - частный критерий: Хи-квадрат, разности между факт...	Семантический резонанс зна...	101	1	0.327	0.980	0.490	28.296	48.060	20.115	0.025	0.885	0.999	0.939
6. INF3 - частный критерий: Хи-квадрат, разности между факт...	Семантический резонанс зна...	101	1	0.327	0.980	0.490	22.550	37.272	14.740	0.025	0.400	0.999	0.939
7. INF4 - частный критерий: ROI (Bayes (Do-Investment), веро...	Семантический резонанс зна...	133	4	0.898	0.990	0.959	20.334	124.515	9.445	0.994	0.994	0.999	0.999
7. INF4 - частный критерий: ROI (Bayes (Do-Investment), веро...	Семантический резонанс зна...	133	4	0.898	1.000	0.941	20.249	2.539	9.554	0.681	1.000	0.999	0.999
8. INF5 - частный критерий: ROI (Bayes (Do-Investment), веро...	Семантический резонанс зна...	133	4	0.898	0.920	0.959	20.334	124.515	9.445	0.994	0.994	0.999	0.999
8. INF5 - частный критерий: ROI (Bayes (Do-Investment), веро...	Семантический резонанс зна...	133	4	0.898	1.000	0.941	20.249	2.539	9.554	0.681	1.000	0.999	0.999
9. INF6 - частный критерий: разности между факт...	Семантический резонанс зна...	122	2	0.282	0.960	0.456	26.861	45.421	24.325	0.239	0.523	0.991	0.991
9. INF6 - частный критерий: разности между факт...	Семантический резонанс зна...	122	1	0.212	0.980	0.549	22.884	14.405	35.687	0.015	0.391	0.999	0.999
10. INF7 - частный критерий: разности между факт...	Семантический резонанс зна...	122	2	0.282	0.960	0.456	26.861	45.421	24.325	0.239	0.523	0.991	0.991
10. INF7 - частный критерий: разности между факт...	Семантический резонанс зна...	122	1	0.212	0.980	0.549	22.884	14.405	35.687	0.015	0.391	0.999	0.999

Рисунок 8. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергера и L1-критерию проф. Е.В.Луценко

Из рисунка 8 мы видим, что в данном интеллектуальном приложении по F-критерию Ван Ризбергена наиболее достоверной является СК-модель INF4 с интегральным критерием «Сумма знаний» ($F=0,939$ при максимуме 1,000), что является довольно хорошим результатом, по критерию L1 проф. Е.В.Луценко та же модель является наиболее достоверной ($L1=0,986$ при максимуме 1,000), что является хорошим результатом.

Это подтверждает наличие и адекватное отражение в СК-модели INF3 сильной причинно-следственной зависимости популярностью песни и ее характеристиками.

На рисунке 9 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений по результатам идентификации характеристик оружия в СК-модели INF3 по данным обучающей выборки:

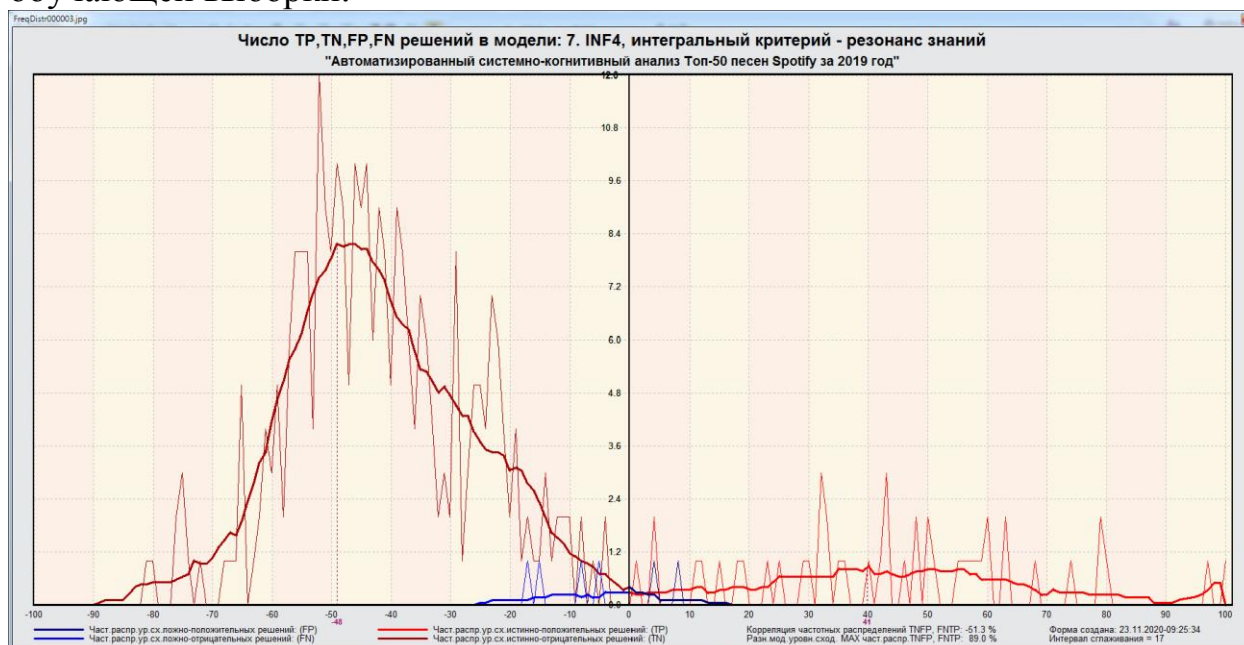


Рисунок 9. Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели Inf4

Рисунок 9 содержит изображения двух частотных распределений, похожих на нормальные, сдвинутых относительно друг друга по фазе.

Левое распределение, большее по амплитуде включает истинно-отрицательные и ложно-положительные решения, а правое, меньшее по амплитуде, включает ложные отрицательные и истинно-положительные решения.

Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу идентификации вида оружия по его характеристикам и другие задачи.

Видно, что для отрицательных решений количество истинных решений всегда значительно превосходит количество ложных решений, причем при уровнях различия больше примерно 5% ложные отрицательные решения вообще практически отсутствуют.

Видно также, что для положительных решений картина более сложная и включает 3 диапазона уровней сходства

1) при уровнях сходства от 0% до 30% количество ложных решений больше числа истинных;

2) при уровнях сходства от 30% до примерно 42% есть и истинные и ложные положительные решения, но число истинных решений больше числа ложных и доля истинных решений возрастает при увеличении уровня сходства;

3) при уровнях сходства выше 42% ложные положительные решения не встречаются.

На рисунке 10 приведен Help по режиму 3.4, в котором описаны меры достоверности моделей, применяемые в системе «Эйдос»:

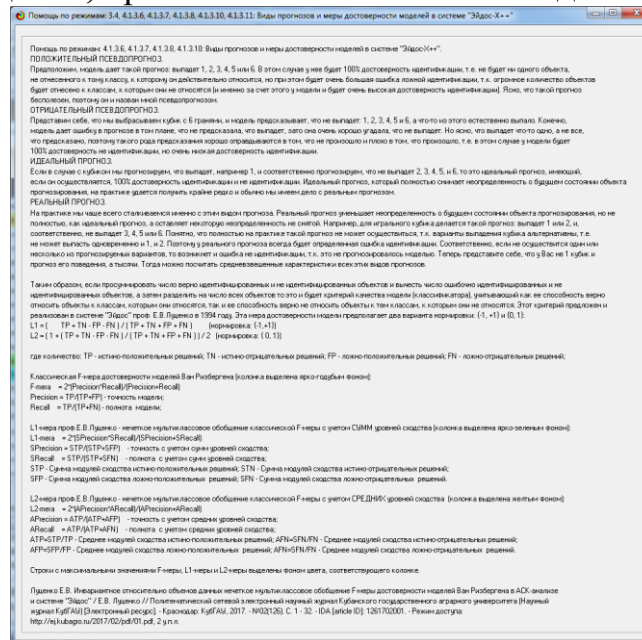


Рисунок 10. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергера и L1- и L2-критериям проф. Е.В.Луценко

Выбор наиболее достоверной модели и присвоение ей статуса текущей

В соответствии со схемой обработки данных, информации и знаний в системе «Эйдос» (рисунок 1), присвоим СК-модели INF3 статус текущей модели. Для этого запустим режим 5.6 с параметрами, приведенными на экранной форме (рисунок 11):

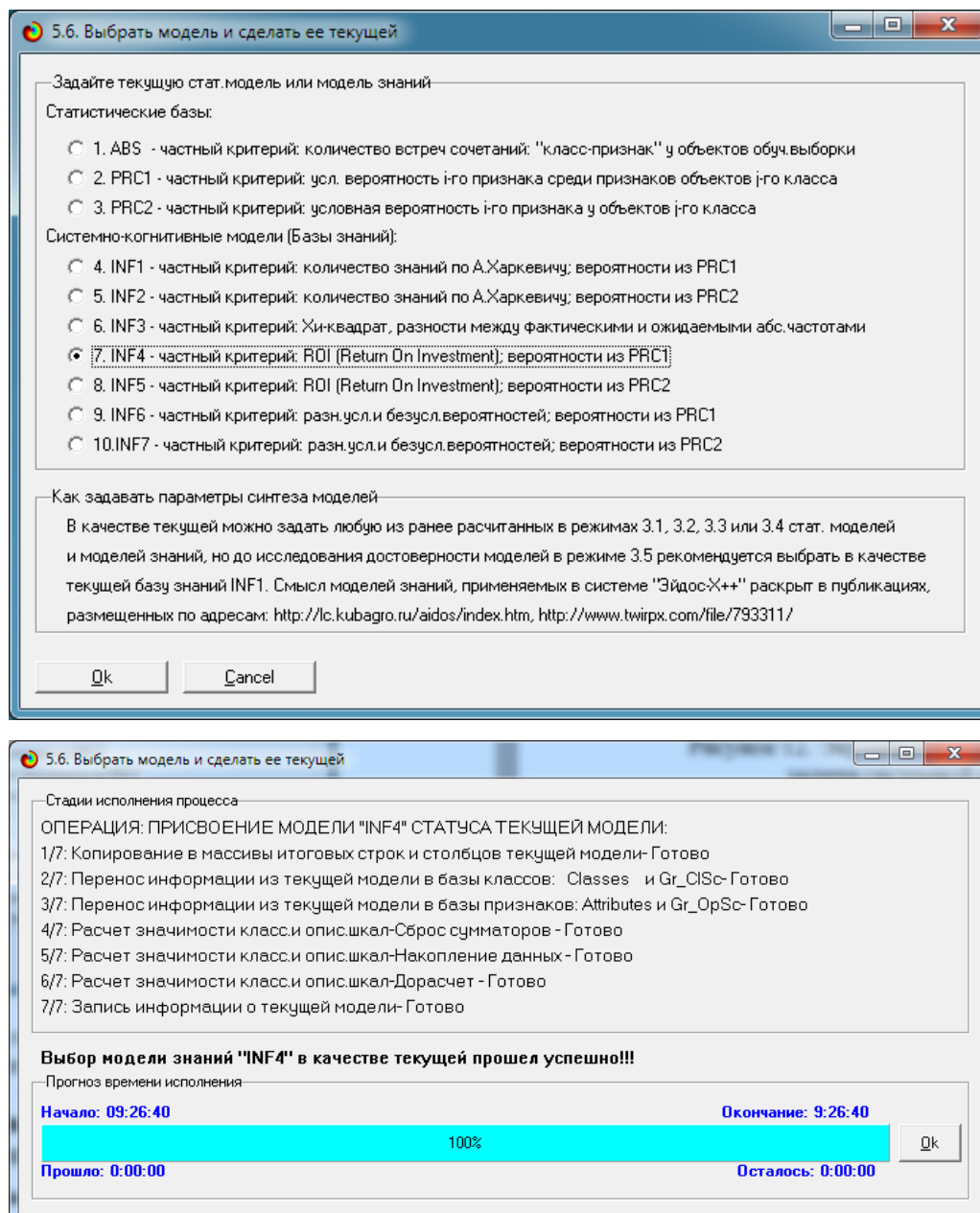


Рисунок 11. Экранные формы придания наиболее достоверной СК-модели Inf3 статуса текущей модели

Задача 4: решение различных задач в наиболее достоверной модели

Подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация)

Решим задачу системной идентификации, т.е. определение класса оружия на основе обучающей выборки в наиболее достоверной СК-модели INF3 на GPU (рисунок 12).

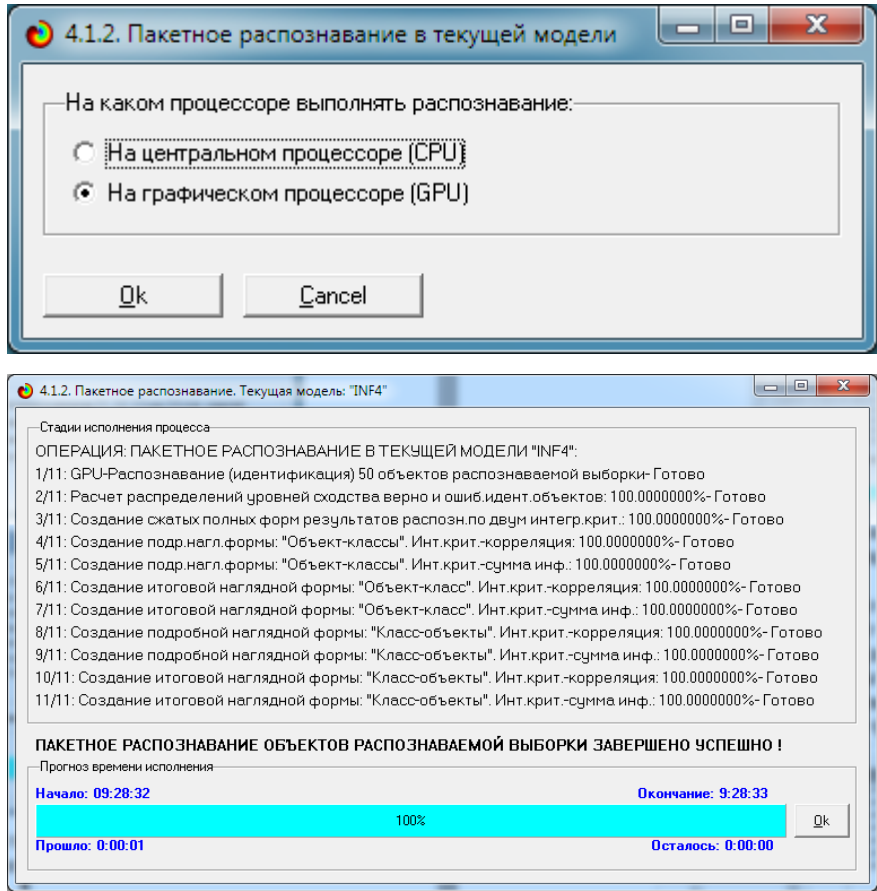
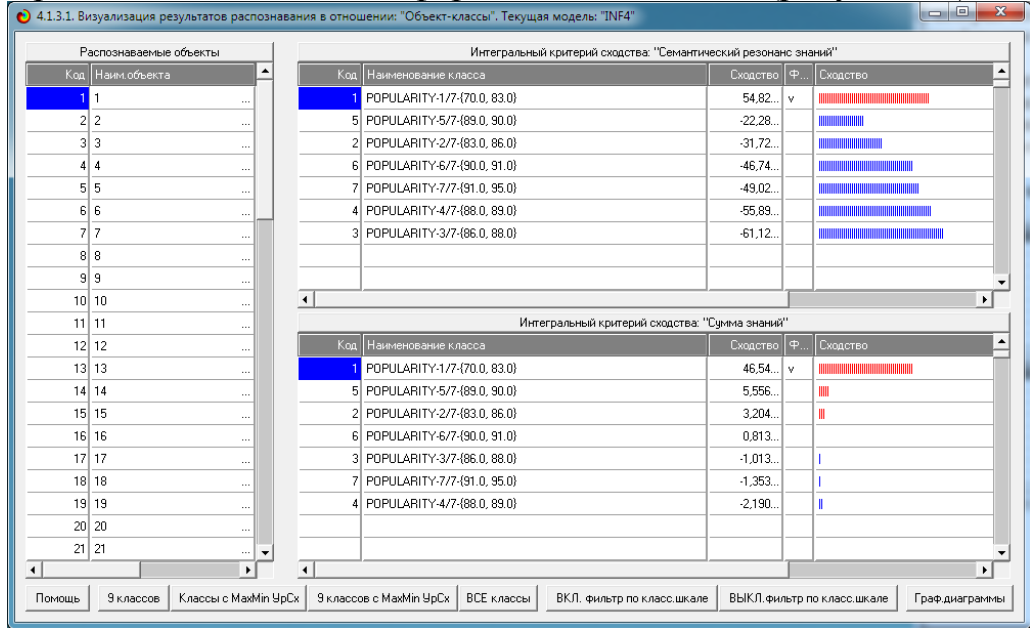


Рисунок 12. Экранные формы, которые отображают процесс решения задачи системной идентификации в текущей модели

Из рисунка 12 видно, что процесс идентификации занял 2 секунды.

Отметим, что 99% этого времени заняла не сама идентификация на GPU, а создание 10 выходных форм на основе результатов этого прогнозирования. Эти формы отражают результаты прогнозирования в различных разрезах и обобщениях:

Приведем две из этих 10 форм: 4.1.3.1 и 4.1.3.2 (рисунок 13).



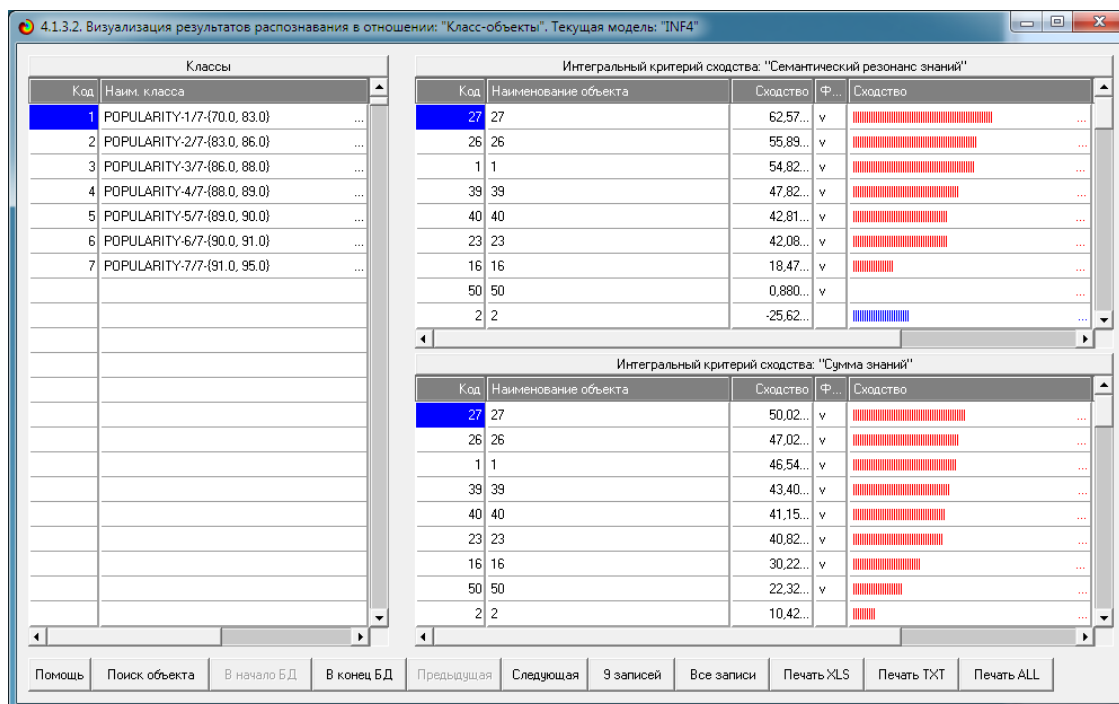


Рисунок 13. Выходные формы по результатам идентификации вида оружия по его характеристикам

Символ «✓» стоит против тех результатов идентификации, которые подтвердились на опыте, т.е. соответствуют факту. Из рисунка 13 видно, что результаты идентификации являются отличными, естественно при учете информации из рисунка 9 о том, что достоверные прогнозы в данной модели имеют уровень сходства выше 42%, т.е., по сути, результаты с более низки уровнем сходства надо просто игнорировать.

Подзадача 4.2. Поддержка принятия решений (SWOT-анализ)

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 задаче, решаемой в данной работе, SWOT-анализ показывает степень влияния различных значений характеристик оружия на его вид.

В системе «Эйдос» в режиме 4.4.8 поддерживается решение этой задачи. При этом **выявляется система детерминации заданного класса**, т.е. система значений факторов, обуславливающих переход объекта моделирования и управления в состояние, соответствующее данному классу, а также препятствующих этому переходу.

На рисунках 14 приведены SWOT-диаграммы наглядно отражающие силу и направление влияния различных значений характеристик песен на ее популярность.

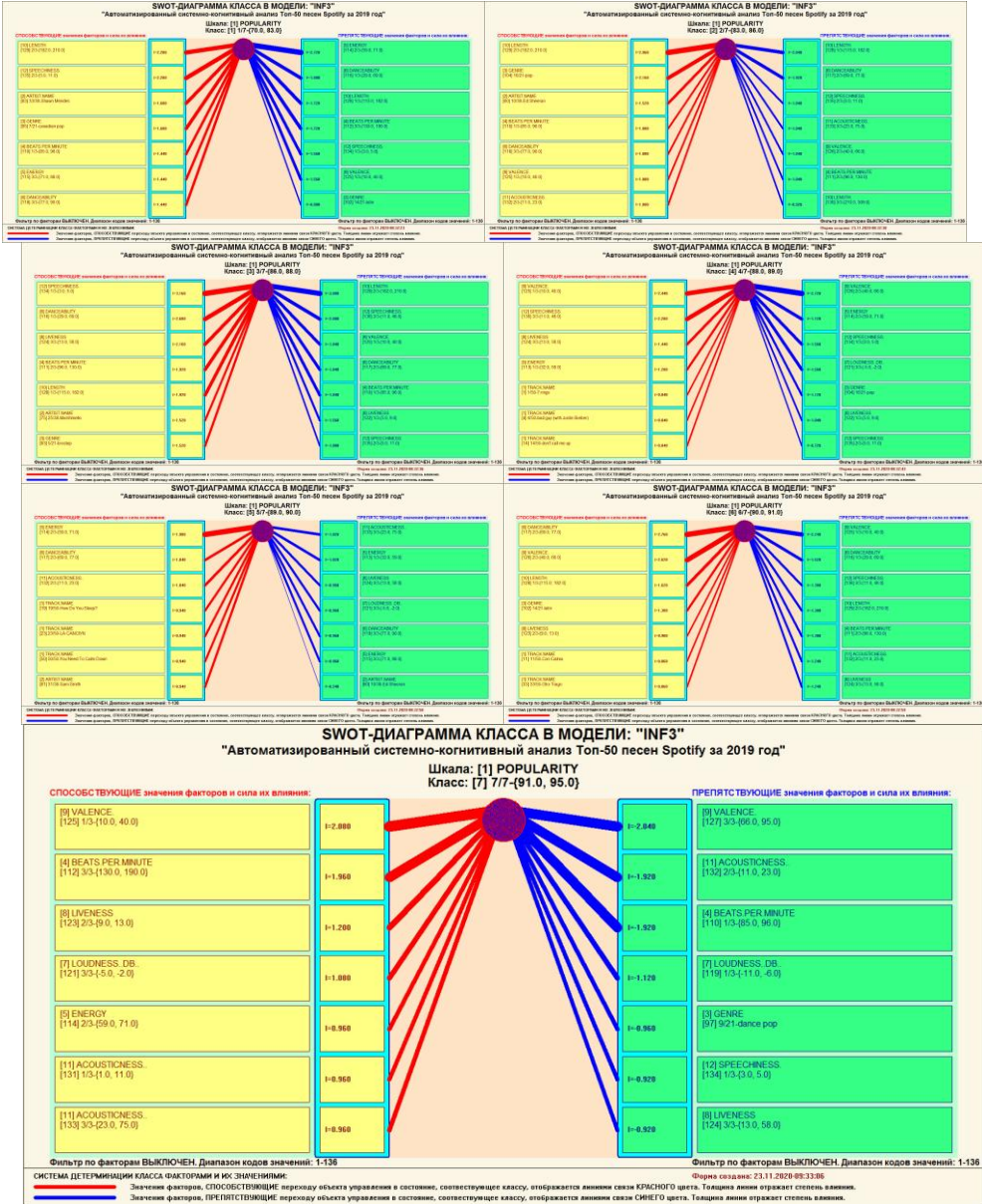


Рисунок 14. SWOT-диаграммы детерминации характеристик оружия

Эти SWOT-диаграммы наглядно отражают силу и направление влияния различных значений характеристик песен на ее популярность.

Отметим также, что система «Эйдос» обеспечивала решение этой задачи *всегда*, т.е. даже в самых ранних DOS-версиях и в реализациях системы «Эйдос» на других языках и типах компьютеров. Например, первый акт внедрения системы «Эйдос», где об этом упоминается в явном виде, датируется 1987 годом, а первый подобный расчет относится к 1981 году.

УТВЕРЖДАЮ
Заведующий Краснодарским
сектором ИСИ АН СССР, к.ф.н.
А.А. Хагуров
19.05.1987г.



УТВЕРЖДАЮ
Директор Северо-Кавказского филиала
ВНИИ "АИУС-агроресурсы", к.э.н.
Э.М. Трахов
19.05.1987г.

А К Т

Настоящий акт составлен комиссией в составе: Кириченко М.М., Ляшко Г.А., Самсонов Г.А., Коренец В.И., Луценко Е.В. в том, что в соответствии с договором о научно-техническом сотрудничестве между Северо-Кавказским филиалом ВНИИ "АИУС-агроресурсы" и Краснодарским сектором Института социологических исследований АН СССР Северо-Кавказским филиалом ВНИИ "АИУС-агроресурсы" выполнены следующие работы:

- осуществлена постановка задачи: "Обработка на ЭВМ социологических анкет Крайагропрома";
- разработаны математическая модель и программное обеспечение подсистемы распознавания образов, позволяющие решать данную задачу в среде персональной технологической системы ВЕГА-М;
- на профессиональной персональной ЭВМ "Искра-226" осуществлены расчёты по задаче в объёме:

Входная информация составила 425 анкет по 9-ти предприятиям. Выходная информация - 4 вида выходных форм объёмом 90 листов формата А3 и 20 листов формата А4 содержит:

- процентное распределение ответов в разрезе по социальным типам корреспондентов;
- распределение информативностей признаков (в битах) для распознавания социальных типов корреспондентов;
- позитивные и негативные информационные портреты 30-ти социальных типов на языке 212 признаков;
- обобщённая характеристика информативности признаков для выбора такого минимального набора признаков, который содержит максимум информации о распознаваемых объектах (оптимизация анкет).

Работы выполнены на высоком научно-методическом уровне и в срок.

От ИСИ АН СССР:

Мл.научный сотрудник

Кириченко М.М.
19.05.1987г.

Мл.научный сотрудник

Ляшко Г.А.
19.05.1987г.

От СКФ ВНИИ "АИУС-агроресурсы":

Зав.отделом аэрокосмических и
тематических изысканий №4, к.э.н.
Самсонов Г.А.
19.05.1987г.

Главный конструктор проекта
Коренец В.И.
19.05.87. 1987г.

Главный конструктор проекта
Луценко Е.В.
19.05.87. 1987г.

Но тогда SWOT-диаграммы назывались позитивным и негативным информационными портретами классов.

Информация о системе значений факторов, обуславливающих переход объекта моделирования в различные будущие состояния, соответствующие классам, может быть приведена не только в диаграммах, показанных на рисунках 14, но и во многих других табличных и графических формах, которые в данной работе не приводятся только из-за ограниченности ее объема. В частности в этих формах может быть выведена значительно более полная информация (в т. ч. вообще вся имеющая в модели). Подобная подробная информация содержится в базах данных, расположенных по пути: \Aidos-X\AID_DATA\A0000003\System\SWOTCls####Inf3.DBF, где: «####» - код класса с ведущими нулями. Эти базы открываются в MS Excel.

В заключение отметим, что 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 чаще всего является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают неформализуемым путем (интуитивно), на основе своего опыта и профессиональной 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 более 30 лет, но, к сожалению, она сравнительно малоизвестна – это интеллектуальная система «Эйдос».

Подзадача 4.3. 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 «Эйдос» 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.

4.3.1. Когнитивные диаграммы классов

Эти диаграммы отражают сходство/различие классов. Мы получаем их в режимах 4.2.2.1 и 4.2.2.2 (рисунок 16).

Отметим также, что на когнитивной диаграмме, приведенной на рисунке 16, показаны **количественные** оценки сходства/различия топ 50 песен в Spotify за 2019 год по связанным с ними значениям их характеристик. 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.

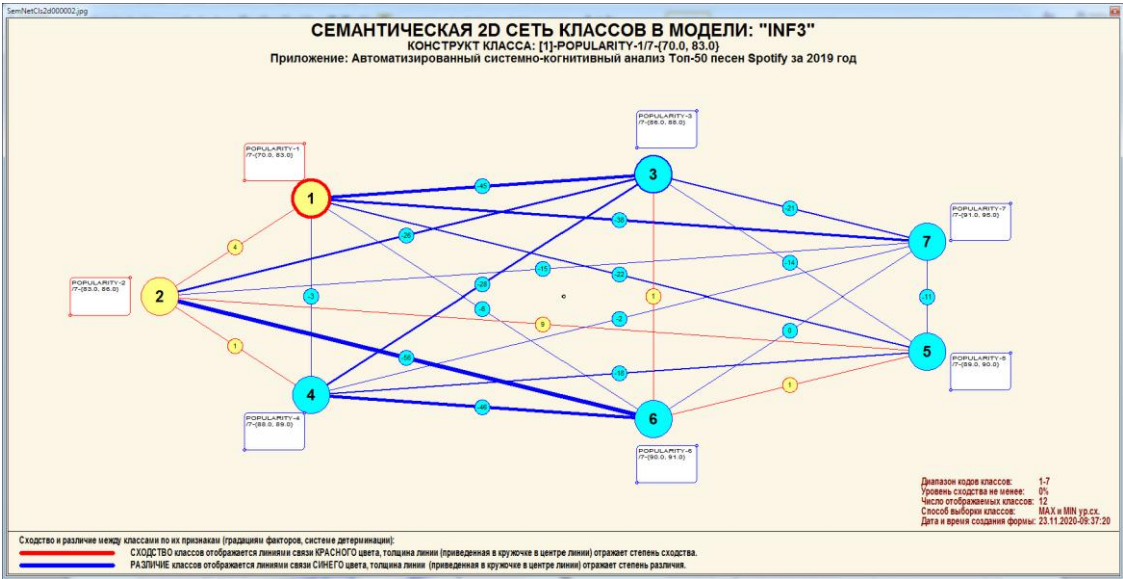
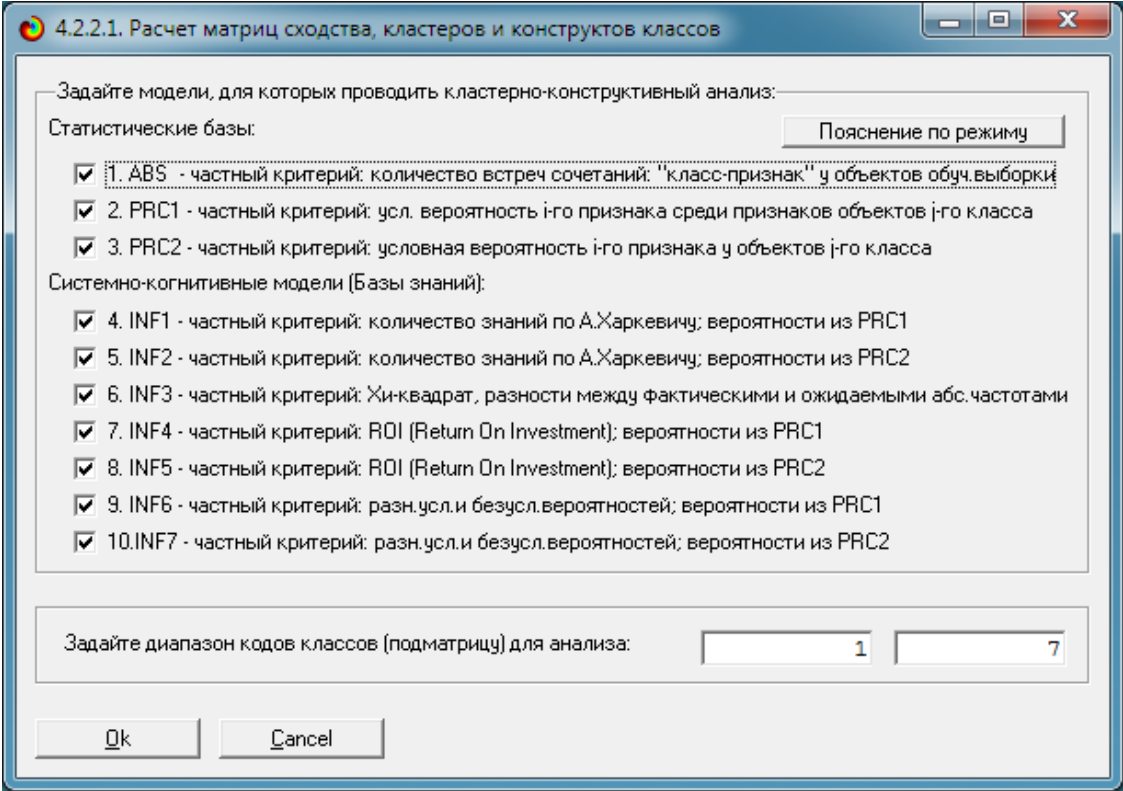


Рисунок 15. Когнитивная диаграмма классов, отражающая сходства/различия песен по связанным с ними значениям их характеристик.

В системе «Эйдос» есть возможность управлять параметрами формирования и вывода изображения, приведенного на рисунке 16. Для этого используется диалоговое окно, приведенное на рисунке 17.

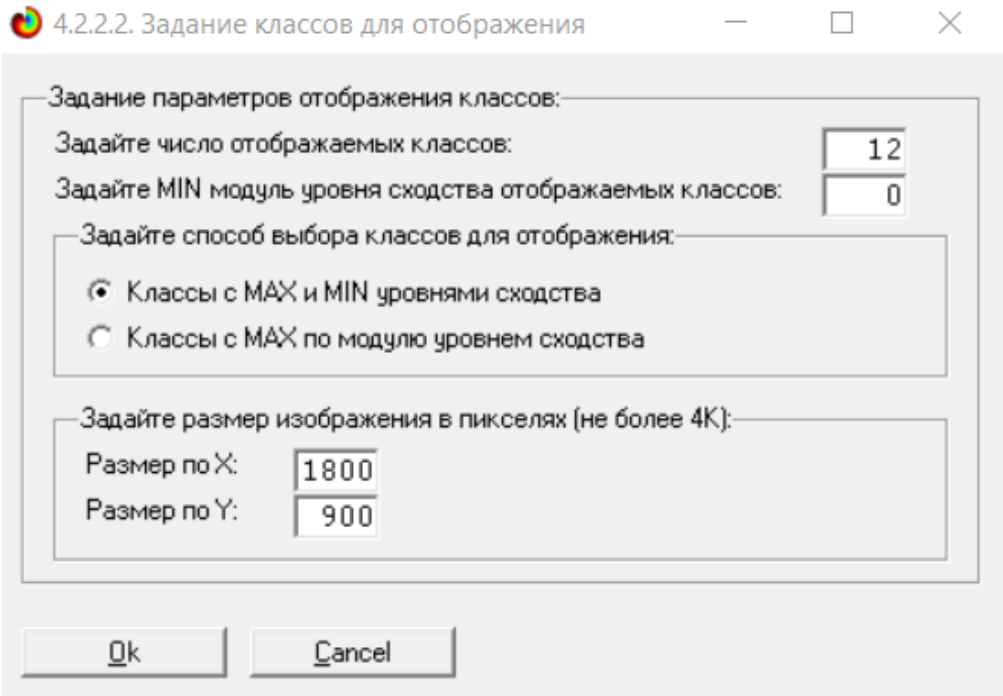


Рисунок 16. Диалоговое окно управления параметрами формирования и вывода изображения когнитивной диаграммы классов

4.3.2. Агломеративная когнитивная кластеризация классов

Информация о сходстве/различии классов, содержащаяся в матрице сходства, может быть визуализирована не только в форме, когнитивных диаграмм, пример которой приведен на рисунке 16, но и в форме агломеративных дендрограмм, полученных в результате *когнитивной кластеризации* (рисунок 18):

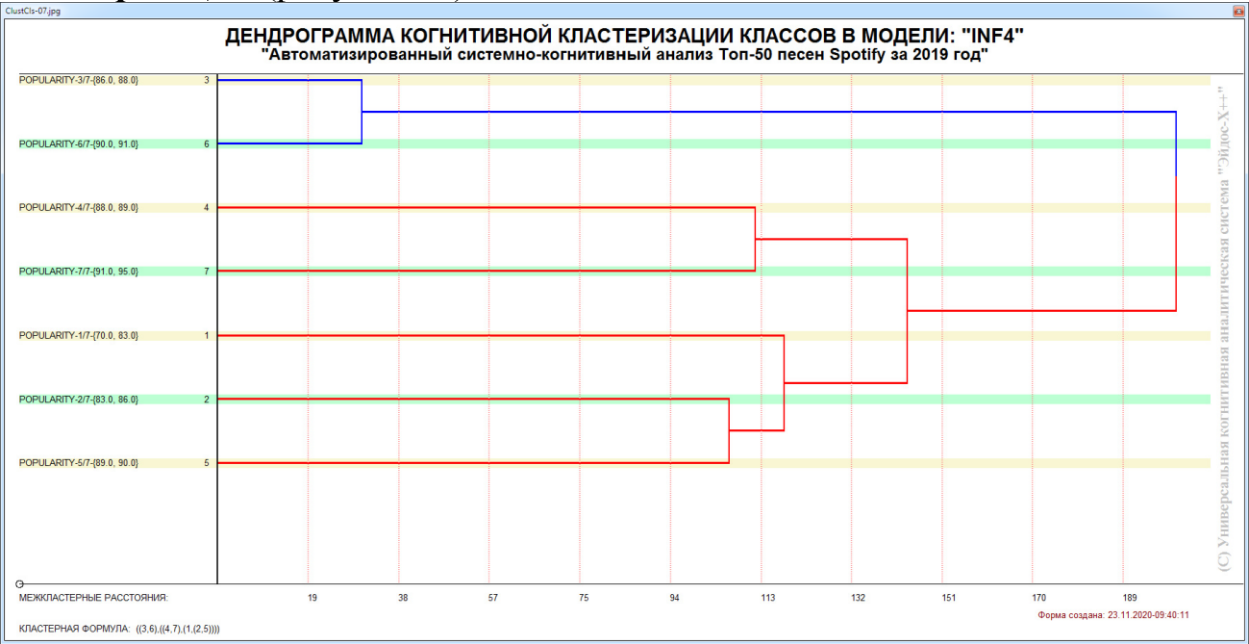


Рисунок 17. Дендрограмма когнитивной агломеративной кластеризации, отражающая сходство/различие песен

Из рисунков 16 и 18 мы видим, что некоторые песни сходны по детерминирующей их системе значений характеристик, и, следовательно, корректно ставить задачу их одновременного достижения, а другие по системе значений этих параметров сильно отличаются, и, следовательно, являются взаимоисключающими, т.е. альтернативными и цель их одновременного достижения является некорректной и недостижимой, т.к. для достижения одного из альтернативных результатов необходимы одни значения характеристик, а для достижения другого – совершенно другие, которые не могут наблюдаться одновременно с первыми.

Из дендрограммы когнитивной агломеративной кластеризации классов, приведенной на рисунке 18, мы видим также, что все классы образуют два противоположных кластера, являющихся полюсами конструкта, по системе значений обуславливающих значения параметров их характеристик.

На рисунке 19 мы видим график изменения межкластерных расстояний:

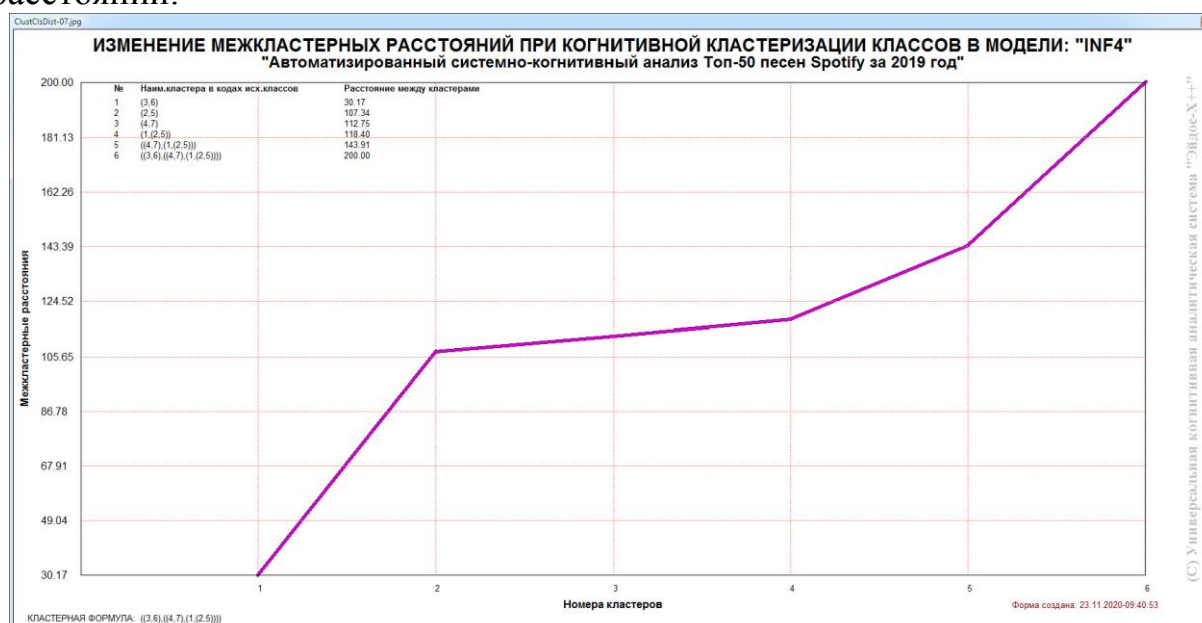
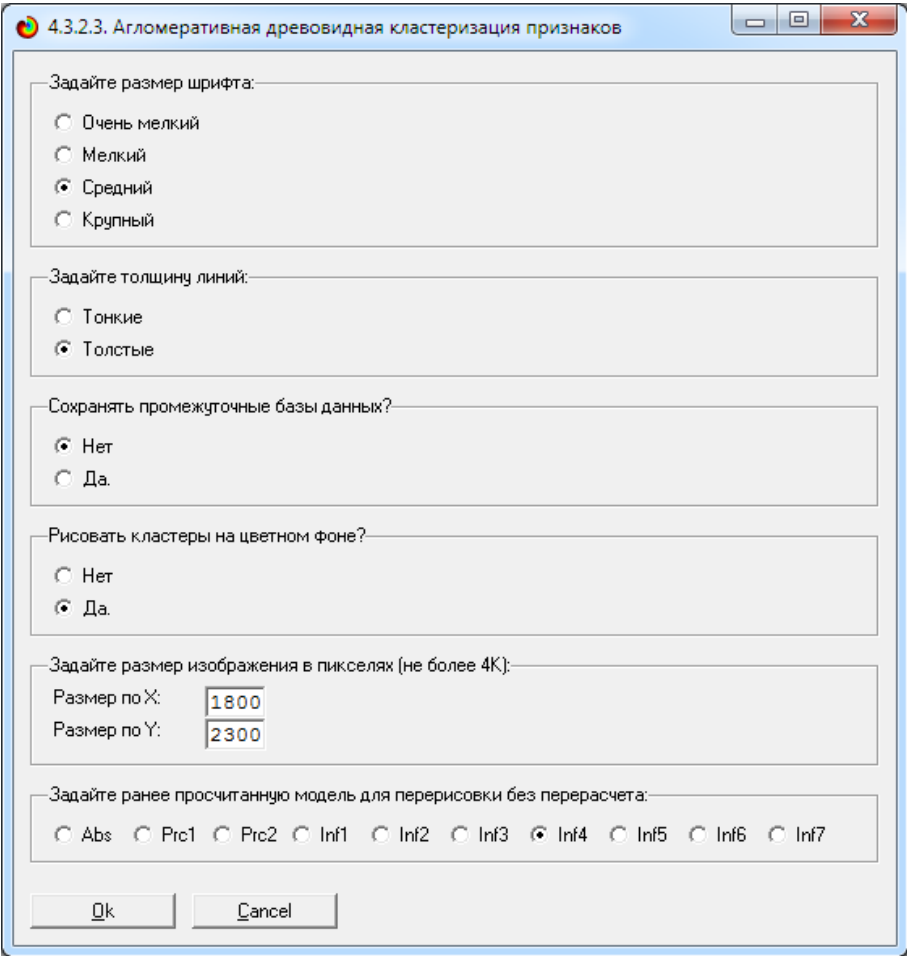
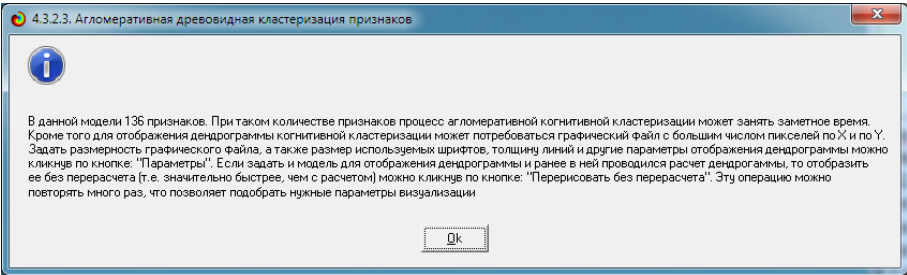
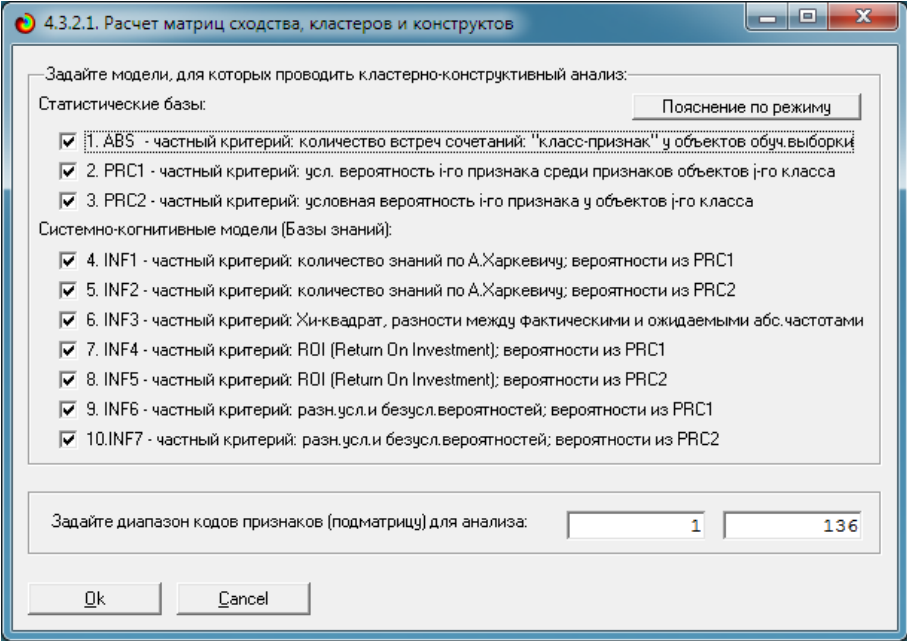


Рисунок 18. График изменения межкластерных расстояний

4.3.3. Агломеративная когнитивная кластеризация значений факторов

На рисунке 22 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, пример которой приведен на рисунке 20.



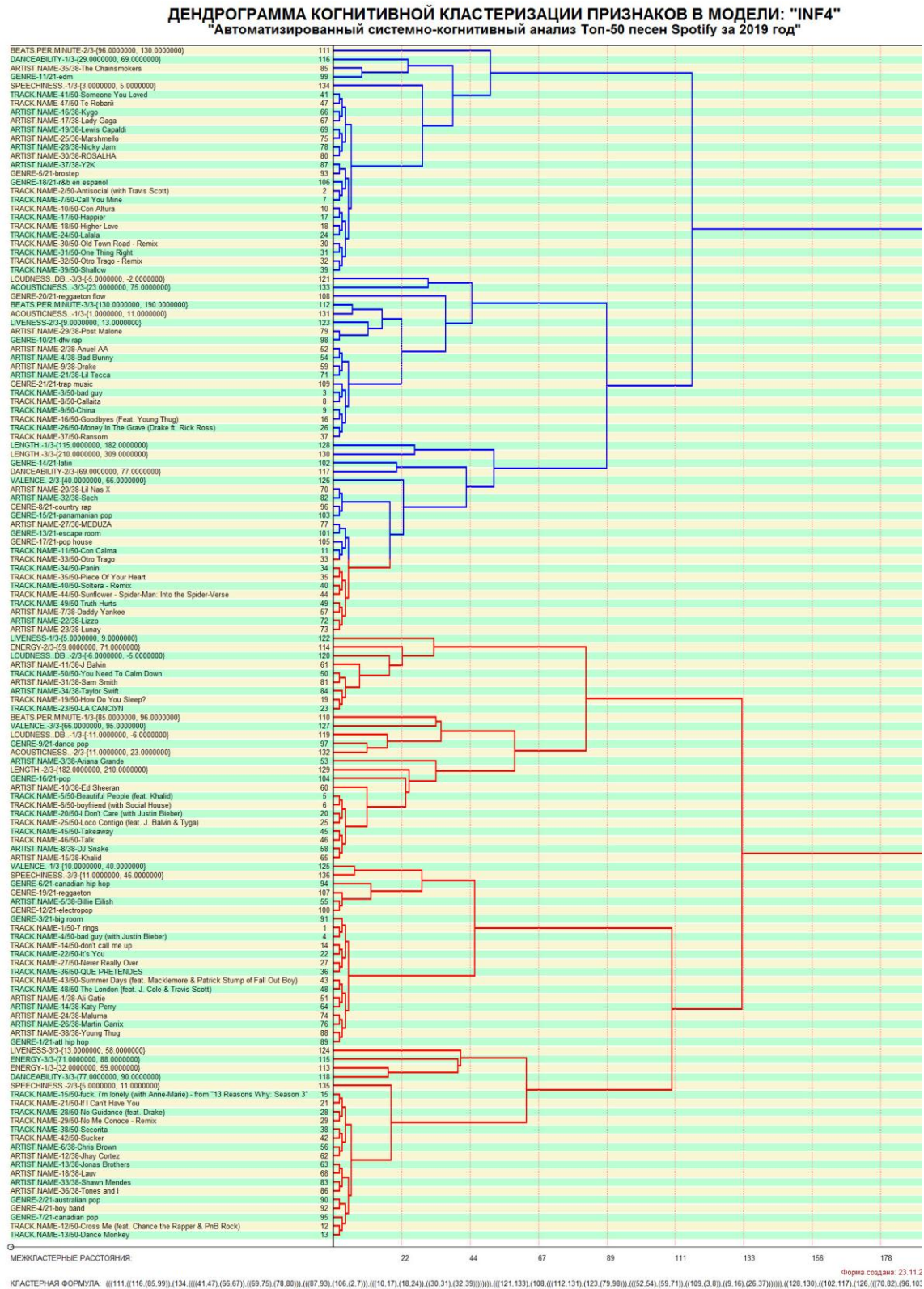


Рисунок 19. Дендрограмма агломеративной когнитивной кластеризации признаков

Из дендрограммы на рисунке 22 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, т.е. по содержащейся в них информации о классе оружия. *Значения факторов на*

Рисунок 210. График изменения межкластерных расстояний при когнитивной кластеризации значений факторов

4.3.4. Нелокальные нейроны и нелокальные нейронные сети

На рисунке 21 приведён пример нелокального нейрона, а на рисунке 22 - фрагмент одного слоя нелокальной нейронной сети:

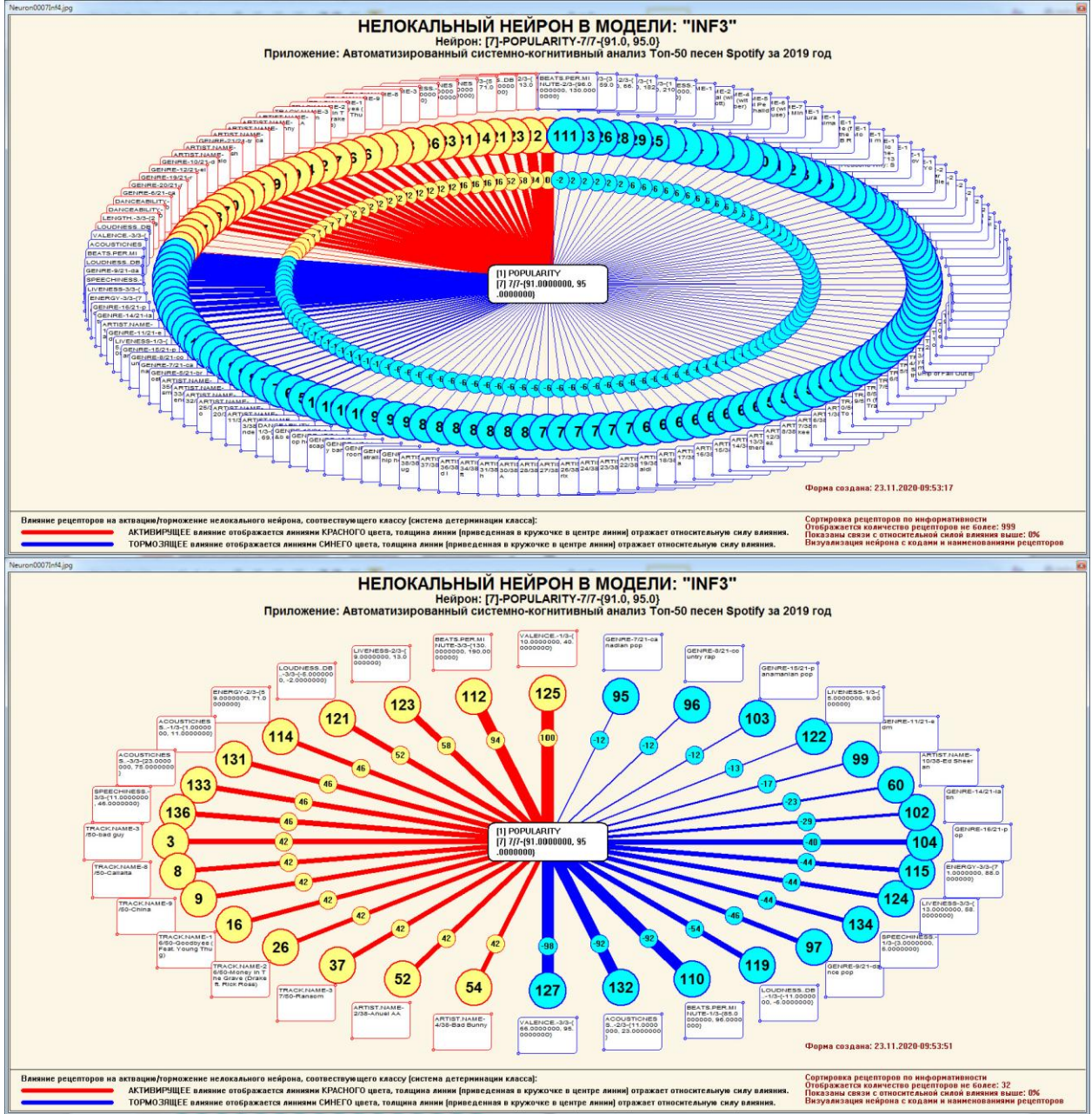


Рисунок 11. Пример нелокального нейрона, отражающего силу и направление влияния значений характеристик песни на ее популярность

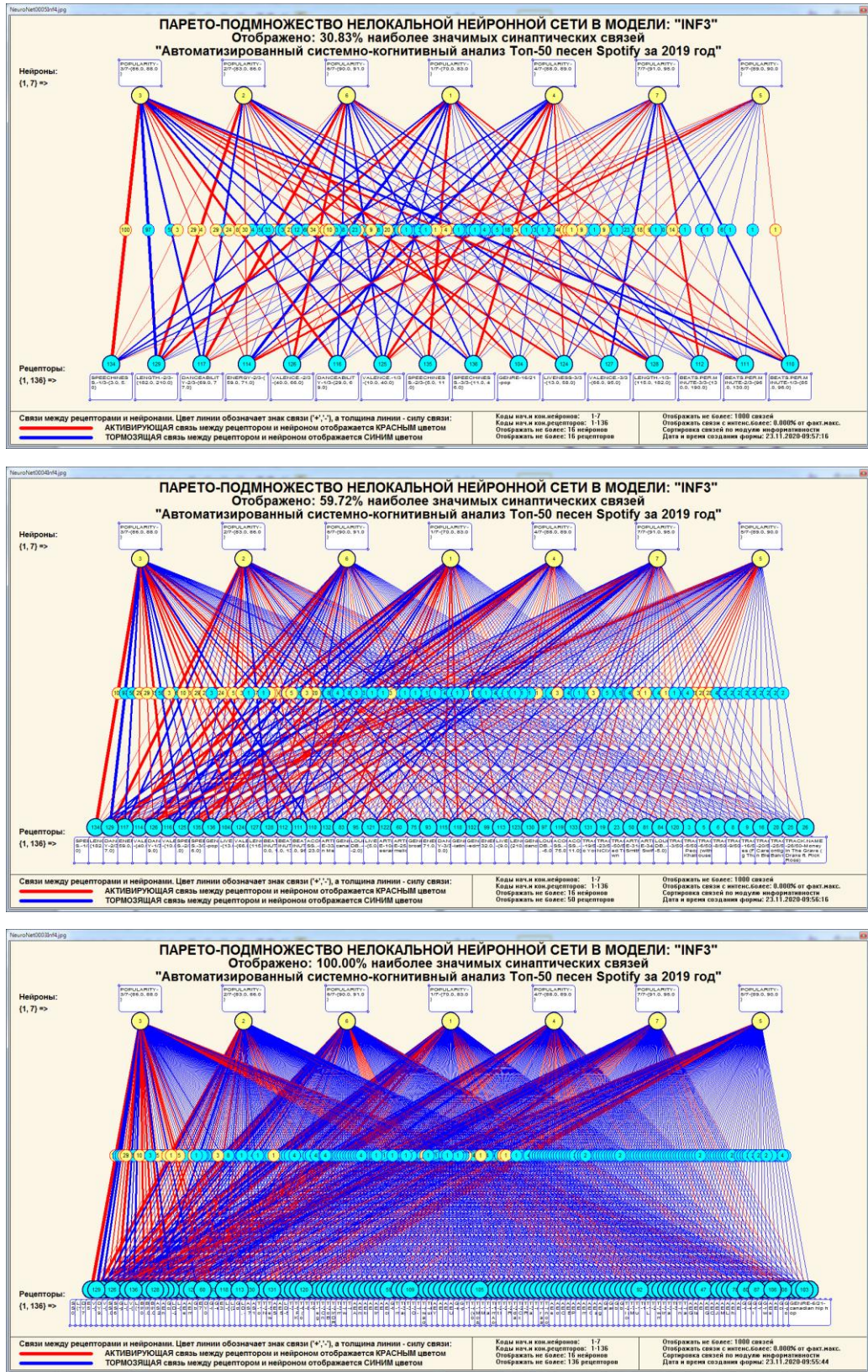


Рисунок 22. Один слой нелокальной нейронной сети, отражающий силу и направление влияния значений характеристик песни на ее популярность (фрагмент 49% сети)

В приведенном фрагменте слоя нейронной сети нейроны соответствуют популярности песен, а рецепторы – их характеристикам. 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.

Модель знаний системы «Эйдос» относится к **нечетким декларативным** гибридным моделям и объединяет в себе некоторые особенности нейросетевой и фреймовой моделей представления знаний. Классы в этой модели соответствуют нейронам и фреймам, а признаки рецепторам и шпациям (описательные шкалы – слотам).

От фреймовой модели представления знаний модель системы «Эйдос» отличается своей эффективной и простой программной реализацией, полученной за счет того, что разные фреймы отличаются друг от друга не набором слотов и шпаций, а лишь информацией в них. Поэтому в системе «Эйдос» при увеличении числа фреймов само количество баз данных не увеличивается, а увеличивается лишь их размерность.

От нейросетевой модели представления знаний модель системы «Эйдос» отличается тем, что:

- 1) весовые коэффициенты на рецепторах не подбираются итерационным методом обратного распространения ошибки, а считаются прямым счетом на основе хорошо теоретически обоснованной модели, основанной на теории информации (это напоминает байесовские сети);
- 2) весовые коэффициенты имеют хорошо теоретически обоснованную содержательную интерпретацию, основанную на теории информации;
- 3) нейросеть является нелокальной, как сейчас говорят «полносвязной».

4.3.5. 3d-интегральные когнитивные карты

На рисунке 26 приведен фрагмент 3d-интегральной когнитивной карты, отражающий фрагмент около 49% СК-модели Inf3.

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивных диаграмм классов и значений факторов, отображенных соответственно на рисунках 16 и 20, и одного слоя нейронной сети, приведенного на рисунке 25.



Рисунок 23. 3d-интегральная когнитивная карта в СК-модели Inf3

4.3.6. Когнитивные функции

Вместо описания того, что представляют собой когнитивные функции, приведем help соответствующего режима системы «Эйдос» (рисунок 24).

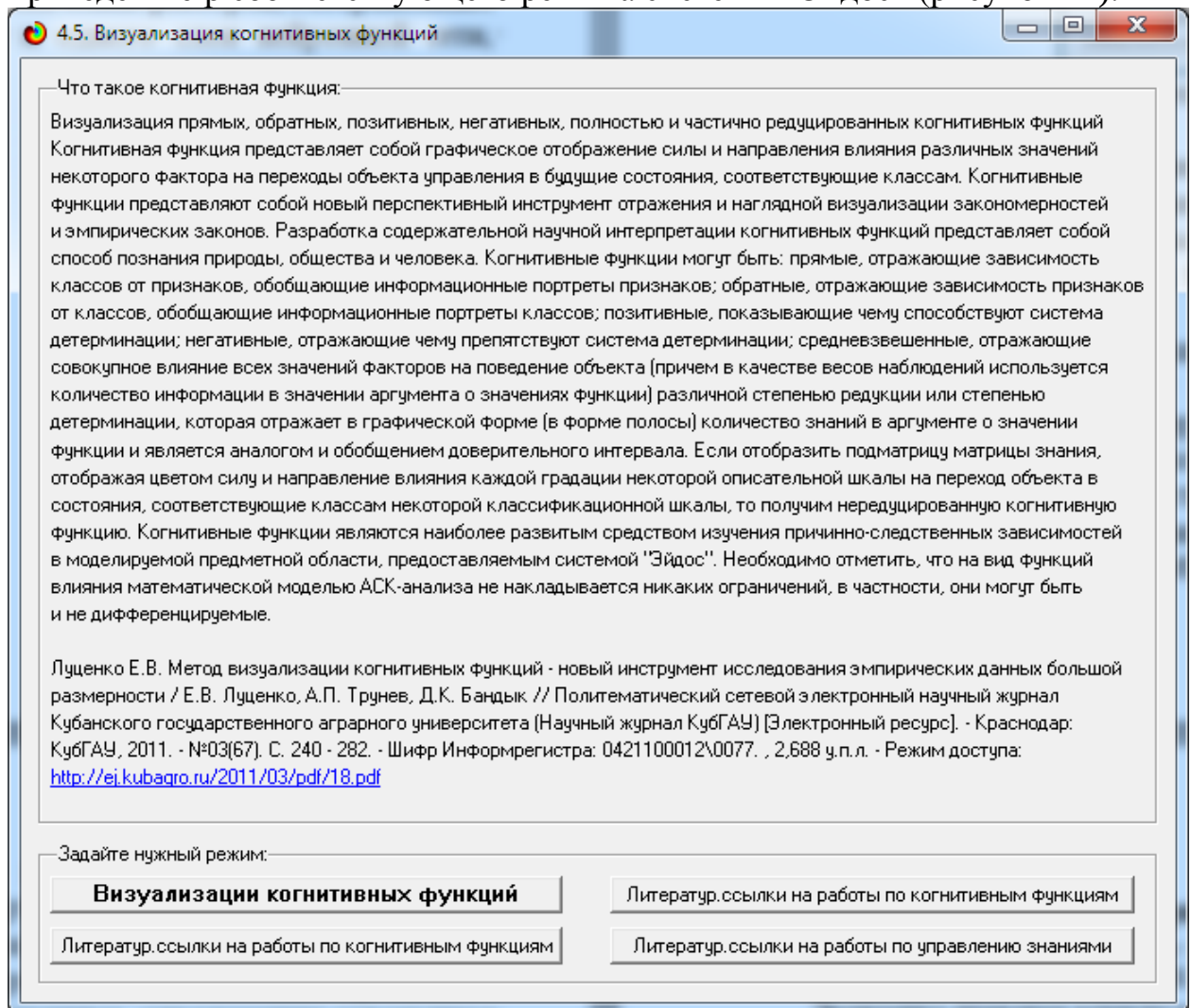


Рисунок 12. Help режима визуализации когнитивных функций

Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора (признаков) на переходы объекта управления в будущие состояния, соответствующие классам. Классы являются градациями классификационных шкал.

Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации эмпирических закономерностей и эмпирических законов. Разработка содержательной научной интерпретации когнитивных функций представляет собой способ познания природы, общества и человека.

Когнитивные функции могут быть: прямые, отражающие зависимость классов от признаков, обобщающие информационные портреты признаков; обратные, отражающие зависимость признаков от классов, обобщающие информационные портреты классов; позитивные, показывающие чему способствуют система детерминации (обозначены белой линией); негативные, отражающие чему препятствуют система детерминации (обозначены черной линией); средневзвешенные, отражающие совокупное влияние всех значений факторов на поведение объекта (причем в качестве весов наблюдений используется количество информации в значении аргумента о значениях функции) различной степенью редукции или степенью детерминации, которая отражает в графической форме (в форме полосы разной толщины) количество знаний в аргументе о значении функции и является аналогом и обобщением доверительного интервала.

Если отобразить подматрицу матрицы знания, отображая цветом силу и направление влияния каждой градации некоторой описательной шкалы на переход объекта в состояния, соответствующие классам некоторой классификационной шкалы, то получим нередуцированную когнитивную функцию.

Когнитивные функции являются наиболее развитым средством изучения причинно-следственных зависимостей в моделируемой предметной области, предоставляемым системой "Эйдос".

Необходимо отметить, что *на вид функций влияния математической моделью АСК-анализа не накладывается никаких ограничений*, в частности, они могут быть и не дифференцируемые.

На рисунках 25 приведены когнитивные функции, наглядно отражающие силу и направление влияния значений (т.е. степени выраженности) различных характеристик песни на ее популярность.

Ниже приведен рисунок 13. Примеры когнитивных функций в СК-модели INF3, отражающих силу и направление влияния значений характеристик песни на ее популярность:

7. Выводы

Как показывает анализ результатов численного эксперимента предложенное и реализованное в системе «Эйдос» решение поставленных задач является вполне эффективным, что позволяет обоснованно утверждать, что цель работы достигнута, поставленная проблема решена.

В результате проделанной работы, с помощью системы «Эйдос» были созданы 3 статистические и 6 системно-когнитивных моделей, в которых непосредственно на основе эмпирических данных сформированы обобщенные образы классов по популярности песен, изучено влияние характеристик топ-50 песен из Spotify за 2019 год на популярность, и, на основе этого, решены задачи идентификации, классификации и исследования моделируемой предметной области путем исследования ее модели.

Со всеми моделями, созданными в данной статье, можно ознакомиться установив облачное Эйдос-приложение №212 в режиме 1.3 системы «Эйдос».

Список литературы

1. Лойко В.И. Подходы к автоматизации процессов управления производством продукции растениеводства / В.И. Лойко, С.А. Курносков, В.В. Ткаченко, Н.А. Ткаченко // Экономико-правовые аспекты реализации стратегии модернизации России: поиск модели эффективного социоэкономического развития: сб. стат. междунар. науч.-практ. конф., Сочи, 5-9 октября 2016 г. – М.: НИИ ЭИП2016. С. 128-132.
2. Луценко Е. В., Лойко В. И., Лаптев В. Н. Системы представления и приобретения знаний : учеб. пособие / Е. В. Луценко, В. И. Лойко, В. Н. Лаптев. – Краснодар: Экоинвест, 2018. – 513 с. ISBN 978-5-94215-415-8. <https://elibrary.ru/item.asp?id=35641755>
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. <http://elibrary.ru/item.asp?id=18632909>
4. Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергера в АСКанализе 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.
5. 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.
6. 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.

7. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системнокогнитивном 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.

8. Луценко Е.В. Моделирование сложных многофакторных нелинейных объектов управления на основе фрагментированных зашумленных эмпирических данных большой размерности в системно-когнитивном анализе и интеллектуальной системе «Эйдос-Х++» / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №07(091). С. 164 – 188. – IDA [article ID]: 0911307012. – Режим доступа: <http://ej.kubagro.ru/2013/07/pdf/12.pdf> 1,562 у.п.л.

9. Луценко Е.В. Открытая масштабируемая интерактивная интеллектуальная online среда для обучения и научных исследований на базе АСК-анализа и системы «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №06(130). С. 1 – 55. – IDA [article ID]: 1301706001. – Режим доступа: <http://ej.kubagro.ru/2017/06/pdf/01.pdf>, 3,438 у.п.л.
http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

10. 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.

11. Луценко Е.В. Системно-когнитивное моделирование влияния агротехнологий на урожайность и качество пшеницы и решение задач прогнозирования, поддержки принятия решений и исследования предметной области / Е.В. Луценко, Е.К. Печурина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2019. – №03(147). С. 62 – 128. – IDA [article ID]: 1471903015. – Режим доступа: <http://ej.kubagro.ru/2019/03/pdf/15.pdf> 4,188 у.п.л.

12. Луценко Е.В., Открытая масштабируемая интерактивная интеллектуальная online среда «Эйдос» («Эйдос-online»). Свид. РосПатента РФ на программу для ЭВМ, Заявка № 2017618053 от 07.08.2017, Гос.рег.№ 2017661153, зарегистр. 04.10.2017. – Режим доступа: <http://lc.kubagro.ru/aidos/2017661153.jpg> 2 у.п.л.

13. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с. ISBN 978-5- 94672-757-0. <http://elibrary.ru/item.asp?id=21358220>