

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

**Факультет заочного обучения**

**РЕЦЕНЗИЯ**

**на курсовой проект (работу)**

Студента (ки) Кораченцова Александра Александрович

курса 2 заочной формы обучения

направления подготовки 09.04.02 «Информационные системы и технологии»

направленность(профиль) «Информационные системы и технологии в науке и управлении» (академическая магистратура)

Наименование темы «Прогнозирование присутствия человека в помещении на основе АСК-анализа показаний датчиков».

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

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

**Оценка качества выполнения курсового проекта (работы)**

№ п/п	Показатель	Оценка соответствия заданию (по 5-и балльной шкале)
1.	Актуальность тематики работы	
2.	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	
3.	Уровень и корректность использования в работе различных методов исследований	
4.	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	
5.	Ясность, четкость, последовательность и обоснованность изложения	
6.	Применение современных технологий обработки информации	
7.	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	
8.	Ответы на вопросы при защите	

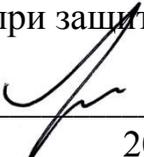
Достоинства работы

Актуальность темы применения современных математических моделей и программного инструментария для проведения АСК-анализа на основе базы данных

Недостатки работы

Отдельные погрешности в оформлении

Итоговая оценка при защите \_\_\_\_\_

Рецензент \_\_\_\_\_  Е.В. Луценко

« \_\_\_\_\_ » \_\_\_\_\_ 201 \_\_\_\_\_

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФГБОУ ВО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ  
ИМЕНИ И.Т. ТРУБИЛИНА»

Факультет заочного обучения  
Кафедра компьютерных технологий и систем

**ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

**к курсовой работе**

по дисциплине: Системы представления и приобретения знаний  
на тему:

**Прогнозирование присутствия человека в помещении на основе АСК-  
анализа показаний датчиков**

выполнил студент группы ИТ31841

**Кораченцов Александр Александрович**

Допущена к защите

Руководитель проекта Луценко Евгений Вениаминович, д.э.н., к.т.н.,  
профессор

(подпись, расшифровка подписи)

Нормоконтролер Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор

(подпись, расшифровка подписи)

Защищена \_\_\_\_\_

Оценка \_\_\_\_\_

\_\_\_\_\_  
(дата)

Члены комиссии \_\_\_\_\_ В.И. Лойко  
\_\_\_\_\_ Е.В. Луценко  
\_\_\_\_\_ В.Н. Лаптев

(подпись, дата, расшифровка подписи)

Краснодар  
2019 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФГБОУ ВО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ  
ИМЕНИ И.Т. ТРУБИЛИНА»

Факультет прикладной информатики  
Кафедра компьютерных технологий и систем

**УТВЕРЖДАЮ:**

**Зав. кафедрой КТС \_\_\_\_\_ В. И. Лойко**

**ЗАДАНИЕ**

**на курсовую работу**

Студенту: ИТ31841 группы 2 курса  
Факультета прикладной информатики  
Специальности: 09.04.02 «Информационные системы и технологии»  
(шифр)

**Кораченцову Александру Александровичу**  
(Ф.И.О.)

Тема проекта: **Прогнозирование присутствия человека в помещении на основе АСК-анализа показаний датчиков**

Содержание задания: Проанализировать современные методы и средства формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами

Объем работы:

а) пояснительная записка к работе \_\_\_\_\_ 47 листов формата А4  
б) графическая часть \_\_\_\_\_ лист формата А4

Рекомендуемая литература: Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос- X++» 2017 г.

Срок выполнения проекта: с “\_\_\_” \_\_\_\_\_ по “\_\_\_” \_\_\_\_\_ 20\_\_ г.

Срок защиты: “\_\_\_” \_\_\_\_\_ 20\_\_ г.

Дата выдачи задания: “\_\_\_” \_\_\_\_\_ 20\_\_ г.

Дата сдачи проекта на кафедру: “\_\_\_” \_\_\_\_\_ 20\_\_ г.

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

(подпись, Ф.И.О., звание, степень)

Задание принял студент

\_\_\_\_\_  
(подпись, дата)

## Реферат

Пояснительная записка курсового проекта (работы) 47 с., 37 рис., 2 табл., 12 источников.

Ключевые слова: УМНЫЙ ДОМ, ДАТЧИК, АВТОМАТИЗАЦИЯ, ОСВЕЩЕНИЕ, ПРИСУТСТВИЕ, ЧЕЛОВЕК, ПРОГНОЗ, ЭЙДОС X, СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ГРАДАЦИИ, КЛАССИФИКАЦИОННЫЕ ШКАЛЫ, ОПИСАТЕЛЬНЫЕ ШКАЛЫ

В курсовом проекте рассматриваются вопросы прогнозирования присутствия человека в помещении на основе АСК-анализа показаний датчиков и системы Эйдос на основе данных репозитория UC1.

## Оглавление

1. Синтез и верификация моделей.....	7
1.1. Описание решения .....	7
1.2. Подготовка и нормализация исходных данных.....	9
1.3. Синтез и верификация статистических и интеллектуальных моделей.....	17
1.4. Верификация статистических и системно-когнитивных моделей....	21
2. Решение задач идентификации в наиболее достоверной модели.....	27
2.1. Выбор наиболее достоверной модели и присвоение ей статуса текущей .....	27
2.2. Прогнозирование (диагностика, классификация, распознавание, идентификация) .....	28
2.3. Визуализация когнитивных функций .....	31
2.4. Нелокальные нейроны и нейронная сеть.....	39
2.5. SWOT и PEST матрицы и диаграммы.....	42
Список литературы .....	46

## Введение

Включение освещения по датчикам движения или присутствия обеспечивает одновременно комфорт и энергосбережение. Свет включается тогда, когда он нужен и выключается, когда не нужен. Всем нам приходилось забывать выключить свет в помещении. Расход электроэнергии на цели освещения может быть заметно снижен достижением оптимальной работы осветительной установки в каждый момент времени. Добиться наиболее полного и точного учета наличия дневного света, равно как и присутствия людей в помещении, можно, применяя средства автоматического управления освещением (СУО) [2].

Однако наиболее перспективным и сложным является автоматический способ управления сценариями по различным событиям. «Умный дом» может работать во многих сценариях. Большинство функций и сцен управляется с пультов управления, компьютеров или с помощью голосового управления, меньшая часть в автоматическом режиме [3].

Управление в автоматическом режиме связано со сложностями, прежде всего с адекватным восприятием системой тех или иных событий. С технической точки зрения данная задача является многокритериальной. Для решения многокритериальных задач применяются элементы искусственного интеллекта.

Используя исторические данные показаний различных датчиков и устройств, таких как, например: влажность, температура, давление, освещенность, содержание CO<sub>2</sub> в воздухе и других, становится возможным прогнозирование присутствия человека в помещении. Для решения задач с использованием искусственного интеллекта будут необходимы обучающие выборки данных.

Создание систем искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий. Так как существует множество альтернатив систем искусственного интеллекта, то возникает необходимость оценки качества

математических моделей этих систем. В данной работе рассмотрено решение задачи по прогнозированию присутствия человека в помещении на основе АСК-анализа показаний датчиков.

Для достижения поставленной цели необходимы свободный доступ к тестовым исходным данным и методика, которая поможет преобразовать эти данные в форму, которая необходима для работы в системе искусственного интеллекта. Удачным выбором является база данных тестовых задач для систем искусственного интеллекта репозитория UCI.

В данной курсовой работе использована база данных «Occupancy Detection Data Set» из банка исходных данных по задачам искусственного интеллекта – репозитория UCI [1]. Для решения задачи используем стандартные возможности Microsoft Office Excel, а также систему искусственного интеллекта «Aidos-X++».

## **1. Синтез и верификация моделей**

### **1.1. Описание решения**

Возможность точного определения присутствия людей в здании в режиме реального времени открывает новые направления автоматизации, включая интеллектуальное управление системами здания для минимизации энергопотребления и визуализацию здания в режиме реального времени. В научной статье [12] авторы оснастили рабочее место в офисе неоднородным массивом датчиков, цель состояла в том, чтобы использовать датчики в тандеме для создания детектора присутствия в реальном времени. В этой работе использовались Деревья принятия решений для выполнения классификации и изучения взаимосвязи между различными типами датчиков, функциями, полученными из данных датчиков, и присутствием человека в помещении.

Авторы обнаружили, что функция, которая лучше всего отличает присутствие от отсутствия – это среднеквадратичное отклонение данных пассивного инфракрасного датчика движения, рассчитанное за

двухминутный период. При использовании с простым пороговым значением эта функция обнаруживает присутствие с точностью 97,9%. Сочетание нескольких датчиков движения с деревом решений позволило повысить точность до 98,4%. Неожиданно, но добавление других типов датчиков, таких как датчик звука, CO<sub>2</sub> и датчика тока, только ухудшило результаты классификации. Как следствие, деревья решений могут улучшить способность системы обнаруживать присутствие, основанные только на датчиках движения, при этом существует риск переобучения, если объединить несколько типов датчиков [12].

Для достижения поставленной цели в данной работе предлагается применить Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную система «Эйдос».

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») [4].

Поэтому в данной работе мы считаем целесообразным привести в упрощенной форме только этапы АСК-анализа, т.к. они, по сути, представляют собой этапы решения поставленных в работе вопросов:

1. Когнитивная структуризация предметной области и подготовка Excel-файла исходных данных;
2. Формализация предметной области, т.е. автоматизированный ввод в систему Эйдос-X++ исходных данных из Excel-файла с помощью стандартного программного интерфейса системы (разработка

классификационных и описательных шкал и градаций и обучающее выборки);

3. Синтез и верификация 3-х статистических и 7 системно-когнитивных моделей;

4. Определение наиболее достоверной модели и придание ей статуса текущей;

5. Решение задач идентификации, диагностики, классификации и прогнозирования;

6. Решение задач поддержки принятия решений;

7. Исследование моделируемой предметной области путем исследования ее модели.

## **1.2. Подготовка и нормализация исходных данных**

Исходные данные для данной работы взяты без изменения из базы данных «Occupancy Detection Data Set» из банка исходных данных по задачам искусственного интеллекта – репозитория UCI [1]. Архив представляет собой три файла с данными. Фрагмент данных из файла приведен на рисунке (Рисунок 1).

Данные представляют собой показания датчиков температуры (Temperature), влажности (Humidity), освещенности (Light) и концентрации углекислого газа (CO<sub>2</sub>) в помещении в определенный момент времени (date). Проверка присутствия человека (Occupancy) в помещении проводилась путем анализа изображения с камеры видеонаблюдения, с интервалом в 1 минуту.

Описание полей:

- date: дата и время в формате ГГГГ-ММ-ДД ЧЧ:ММ:СС;
- Temperature: температура в градусах Цельсия;
- Relative Humidity: относительная влажность, %;
- Light: освещенность, Люкс;

– CO2: концентрация углекислого газа, parts per million (ppm) – частей на миллион;

– Humidity Ratio: удельная влажность, отношение массы (веса) водяного пара к массе (весу) влажного воздуха в том же объеме;

– Occupancy: присутствие, 0 или 1, 0 – людей в помещении нет, 1 – люди есть.

```
"date","Temperature","Humidity","Light","CO2","HumidityRatio","Occupancy"  
"140","2015-02-02 14:19:00",23.7,26.272,585.2,749.2,0.00476416302416414,1  
"141","2015-02-02 14:19:59",23.718,26.29,578.4,760.4,0.00477266099212519,1  
"142","2015-02-02  
14:21:00",23.73,26.23,572.6666666666667,769.6666666666667,0.00476515255246541,1  
"143","2015-02-02 14:22:00",23.7225,26.125,493.75,774.75,0.00474377335599685,1  
"144","2015-02-02 14:23:00",23.754,26.2,488.6,779,0.00476659399998615,1  
"145","2015-02-02 14:23:59",23.76,26.26,568.6666666666667,790,0.00477933243163454,1  
"146","2015-02-02 14:25:00",23.73,26.29,536.3333333333333,798,0.00477613633274892,1  
"147","2015-02-02 14:25:59",23.754,26.29,509,797,0.00478309370839038,1  
"148","2015-02-02 14:26:59",23.754,26.35,476,803.2,0.00479409399662041,1  
"149","2015-02-02 14:28:00",23.736,26.39,510,809,0.00479618871038935,1  
"150","2015-02-02 14:29:00",23.745,26.445,481.5,815.25,0.00480888622067716,1
```

Рисунок 1 – Фрагмент данных из базы данных «Occupancy Detection Data Set»

Поскольку ввод исходных данных в систему «Эйдос» планируется осуществить с помощью ее универсального программного интерфейса импорта данных из внешних баз данных, который работает с файлами MS Excel, то преобразуем данные из csv-файла в xls-файл. С исходными данными были выполнены следующие преобразования:

– файлы с данными были объединены для получения единой выборки, объем выборки: 12417 строк;

– замена всех запятых на точку с запятой;

– замена точек на запяты;

– в столбце Occupancy произведена замена: 0 – No, 1 – Yes;

– удалено поле date и были добавлены два новые поля: день недели и час

- добавлено поле id с инкрементальной нумерацией;
- столбец Occupancy был перенесен во второй столбец;
- файл сохранен с именем inp\_data.xlsx.

Фрагмент подготовленной и нормализованной выборке приведен на рисунке (Рисунок 2).

	A	B	C	D	E	F	G	H	I	J
1	id	Occupancy	dayofweek	hour	Temperat	Humidity	Light	CO2	HumidityRatio	
2	1	Yes	Tuesday	14	23,7	26,272	585,2	749,2	0,004764	
3	2	Yes	Tuesday	14	23,718	26,29	578,4	760,4	0,004773	
4	3	Yes	Tuesday	14	23,73	26,23	572,6667	769,6667	0,004765	
5	4	Yes	Tuesday	14	23,7225	26,125	493,75	774,75	0,004744	
6	5	Yes	Tuesday	14	23,754	26,2	488,6	779	0,004767	
7	6	Yes	Tuesday	14	23,76	26,26	568,6667	790	0,004779	
8	7	Yes	Tuesday	14	23,73	26,29	536,3333	798	0,004776	
9	8	Yes	Tuesday	14	23,754	26,29	509	797	0,004783	
10	9	Yes	Tuesday	14	23,754	26,35	476	803,2	0,004794	
11	10	Yes	Tuesday	14	23,736	26,39	510	809	0,004796	
12	11	Yes	Tuesday	14	23,745	26,445	481,5	815,25	0,004809	
13	12	Yes	Tuesday	14	23,7	26,56	481,8	824	0,004817	
14	13	Yes	Tuesday	14	23,7	26,6	475,25	832	0,004824	
15	14	Yes	Tuesday	14	23,7	26,7	469	845,3333	0,004842	
16	15	Yes	Tuesday	14	23,7	26,774	464	852,4	0,004856	
17	16	Yes	Tuesday	14	23,7	26,89	464	861	0,004877	
18	17	Yes	Tuesday	14	23,7	26,9725	455	880	0,004892	
19	18	Yes	Tuesday	14	23,6	26,89	454	891	0,004848	
20	19	Yes	Tuesday	14	23,64	26,976	458	897,6	0,004875	
21	20	Yes	Tuesday	14	23,65	27,05	464	900,5	0,004891	
22	21	Yes	Tuesday	14	23,64	27,1	473	908,8	0,004898	
23	22	Yes	Tuesday	14	23,6	27,16	464	918	0,004897	
24	23	Yes	Tuesday	14	23,6	27,236	498,4	925,2	0,00491	

Рисунок 2 – Фрагмент данных обучающей выборки

Полученный файл необходимо разместить в папке «Эйдос»: \idos-X\AID\_DATA\Inp\_data\. Автоматизированная формализация предметной области происходит путем импорта исходных данных из внешних баз данных в систему «Эйдос».

Для загрузки базы исходных данных в систему «Эйдос» необходимо воспользоваться универсальным программным интерфейсом для ввода данных из внешних баз данных табличного вида, т.е. режимом 2.3.2.2.

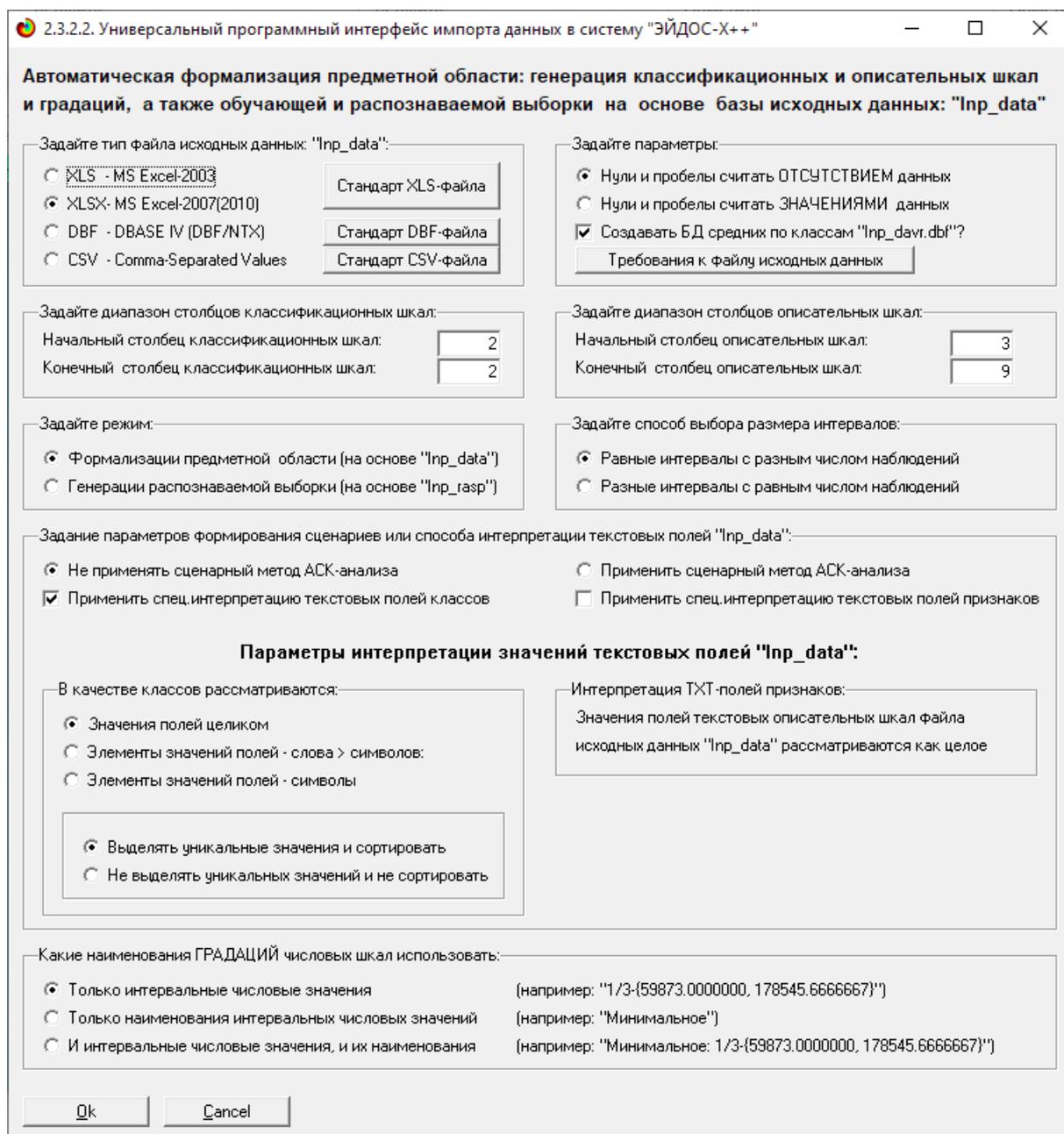


Рисунок 3 – Экранная форма универсального программного интерфейса импорта данных в систему "Эйдос" (режим 2.3.2.2)

В экранной форме, приведенной на рисунке (Рисунок 3), необходимо задать следующие настройки:

– «Задайте тип файла исходных данных «Inp\_data»: «XLSX – MS Excel-2007(2010)»;

– «Задайте диапазон столбцов классификационных шкал»: «Начальный столбец классификационных шкал» – 2;

– «Конечный столбец классификационных шкал» – 2 (второй столбец в таблице);

– «Задайте диапазон столбцов описательных шкал»: «Начальный столбец описательных шкал» – 3, «Конечный столбец описательных шкал» – 9;

– «Задание параметров формирования сценариев или способа интерпретации текстовых полей»: «Не применять сценарный метод АСК-анализа», «Применить спец. интерпретацию текстовых полей классов».

После заполнения необходимых данных, нажать кнопку «Ок». Через некоторое время появляется окно внутреннего калькулятора (Рисунок 4). Экранная форма позволяет оценить, сколько текстовых и числовых классификационных и описательных шкал система обнаружила при заданных параметрах и сколько в них обнаружено градаций.

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных шкал	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	0	0	0,00	6	60	10,00
Текстовые	1	2	2,00	1	7	7,00
ВСЕГО:	1	2	2,00	7	67	9,57

Задайте число интервалов (градаций) в шкале: \_\_\_\_\_

В описательных шкалах:

Пересчитать шкалы и градации      Выйти на создание модели

Рисунок 4 – Экранная форма внутреннего калькулятора режима 2.3.2.2

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.

Если обнаружены шкалы числового типа, то появляется возможность задать количество интервальных числовых значений в этих шкалах. Это делается отдельно для классификационных и описательных шкал, таким образом, число интервальных числовых значений в классификационных и описательных шкалах может отличаться. Если это число изменяется, то необходимо кликнуть полевой кнопке, а затем уже выходить на создание модели. В результате выполнения данного режима формируются классификационные, описательные шкалы, градации и обучающая выборка.

После нажатия на кнопку «Выйти на создание модели» запускается процесс импорта данных из внешней БД «Inp\_data.xlsx» в систему «Эйдос» (Рисунок 5), а также прогноз времени завершения этого процесса.

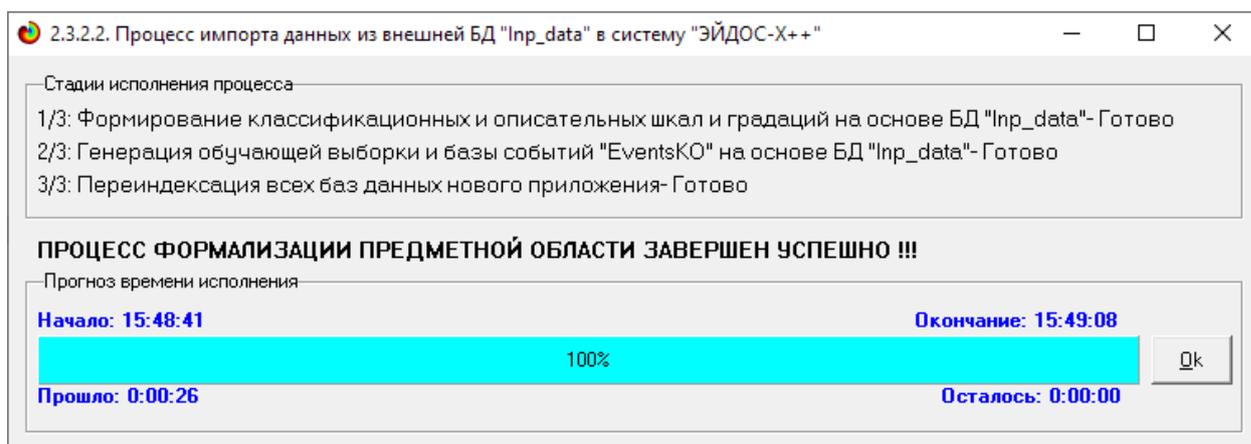


Рисунок 5 – Экранная форма импорта данных режима 2.3.2.2

В результате работы режима сформировано 1 классификационная шкала с шкала с суммарным количеством градаций (классов): 2 (Таблица 1) и 7 описательных шкал с суммарным числом градаций: 67 (

Таблица 2).

С использованием классификационных и описательных шкал и градаций исходные данные (Рисунок 2) были закодированы и в результате получена обучающая выборка. Для просмотра классификационных шкал и

градаций используется режим 2.1, для просмотра описательных шкал и градаций используется режим 2.2

Таблица 1 – Классификационные шкалы и градации

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации
1	OCCUPANCY	1	No
		2	Yes

Таблица 2 – Описательные шкалы и градации

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации
1	DAYOFWEEK	1	Friday
		2	Monday
		3	Saturday
		4	Sunday
		5	Thursday
		6	Tuesday
		7	Wednesday
2	HOUR	8	1/10-{1.0000000,3.2000000}
		9	2/10-{3.2000000,5.4000000}
		10	3/10-{5.4000000,7.6000000}
		11	4/10-{7.6000000,9.8000000}
		12	5/10-{9.8000000,12.0000000}
		13	6/10-{12.0000000,14.2000000}
		14	7/10-{14.2000000,16.4000000}
		15	8/10-{16.4000000,18.6000000}
		16	9/10-{18.6000000,20.8000000}
		17	10/10-{20.8000000,23.0000000}
3	TEMPERATURE	18	1/10-{19.5000000, 19.9908333}
		19	2/10-{19.9908333, 20.4816667}
		20	3/10-{20.4816667, 20.9725000}
		21	4/10-{20.9725000, 21.4633333}
		22	5/10-{21.4633333, 21.9541667}
		23	6/10-{21.9541667, 22.4450000}
		24	7/10-{22.4450000, 22.9358333}
		25	8/10-{22.9358333, 23.4266666}
		26	9/10-{23.4266666, 23.9175000}
		27	10/10-{23.9175000, 24.4083333}
4	HUMIDITY	28	1/10-{21.8650000, 23.6285000}
		29	2/10-{23.6285000, 25.3920000}
		30	3/10-{25.3920000, 27.1555000}
		31	4/10-{27.1555000, 28.9190000}

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации
		32	5/10-{28.9190000, 30.6825000}
		33	6/10-{30.6825000, 32.4460000}
		34	7/10-{32.4460000, 34.2095000}
		35	8/10-{34.2095000, 35.9730000}
		36	9/10-{35.9730000, 37.7365000}
		37	10/10-{37.7365000, 39.5000000}
5	LIGHT	38	1/10-{4.6666667, 173.9250000}
		39	2/10-{173.9250000, 343.1833334}
		40	3/10-{343.1833334, 512.4416667}
		41	4/10-{512.4416667, 681.7000000}
		42	5/10-{681.7000000, 850.9583334}
		43	6/10-{850.9583334, 1020.2166667}
		44	7/10-{1020.2166667, 1189.4750000}
		45	8/10-{1189.4750000, 1358.7333333}
		46	9/10-{1358.7333333, 1527.9916667}
		47	10/10-{1527.9916667, 1697.2500000}
6	CO2	48	1/10-{427.5000000, 592.4000000}
		49	2/10-{592.4000000, 757.3000000}
		50	3/10-{757.3000000, 922.2000000}
		51	4/10-{922.2000000, 1087.1000000}
		52	5/10-{1087.1000000, 1252.0000000}
		53	6/10-{1252.0000000, 1416.9000000}
		54	7/10-{1416.9000000, 1581.8000000}
		55	8/10-{1581.8000000, 1746.7000000}
		56	9/10-{1746.7000000, 1911.6000000}
		57	10/10-{1911.6000000, 2076.5000000}
7	HUMIDITYRATIO	58	1/10-{0.0032748, 0.0035242}
		59	2/10-{0.0035242, 0.0037736}
		60	3/10-{0.0037736, 0.0040229}
		61	4/10-{0.0040229, 0.0042723}
		62	5/10-{0.0042723, 0.0045217}
		63	6/10-{0.0045217, 0.0047711}
		64	7/10-{0.0047711, 0.0050205}
		65	8/10-{0.0050205, 0.0052698}
		66	9/10-{0.0052698, 0.0055192}
		67	10/10-{0.0055192, 0.0057686}

Обучающая выборка, по сути, представляет собой нормализованную с помощью классификационных и описательных шкал и градаций базу исходных данных. Тем самым создаются все необходимые и достаточные предпосылки для выявления силы и направления причинно-следственных

связей между значениями факторов и результатами их совместного системного воздействия (с учетом нелинейности системы [5]). Это делает исходные данные готовыми для обработки в программной системе и выполнения следующего этапа АСК-анализа: синтеза и верификации модели.

### 1.3. Синтез и верификация статистических и интеллектуальных моделей

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) осуществляется в режиме 3.5 системы «Эйдос» (Рисунок 6).

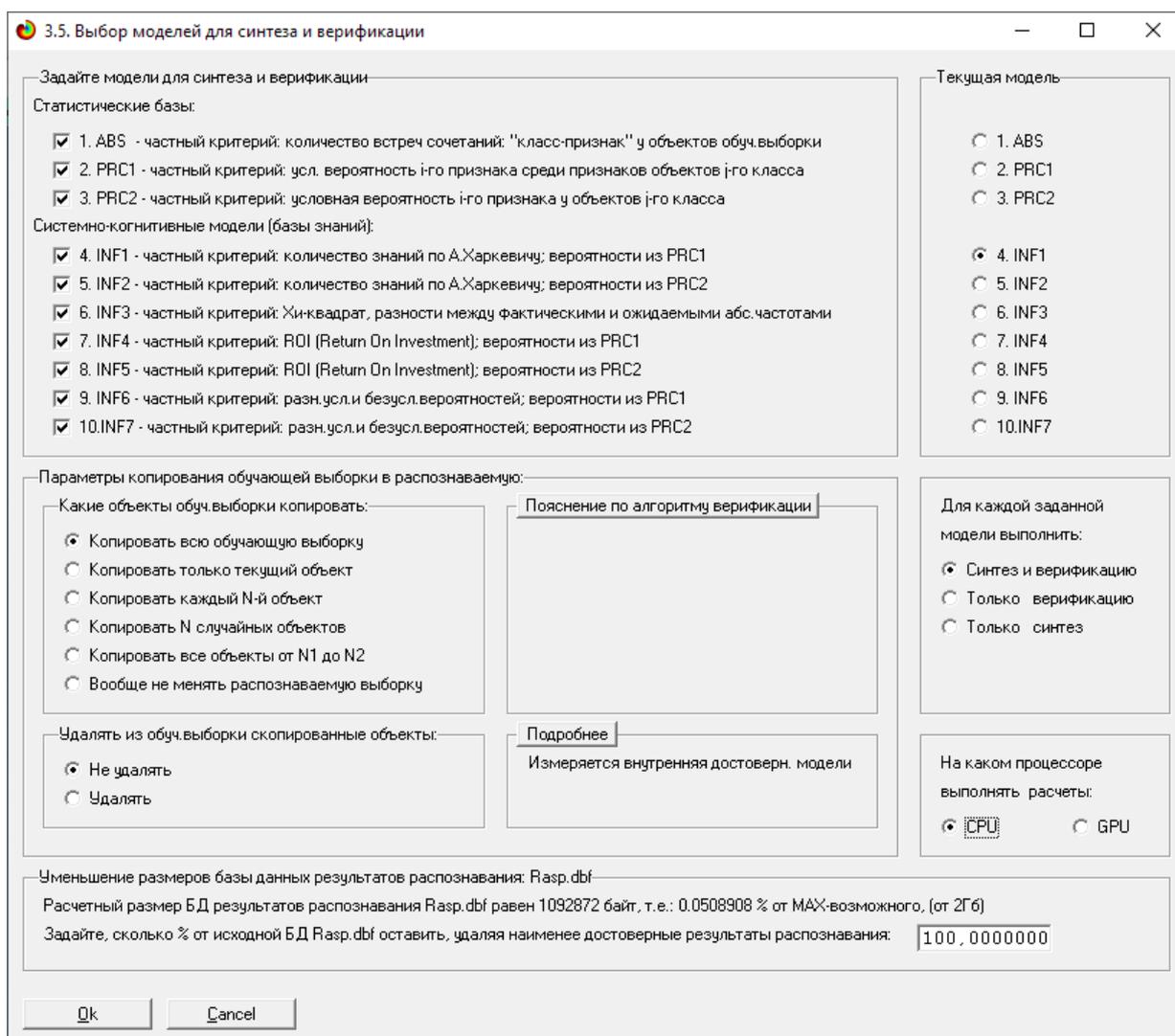


Рисунок 6 – Экранная форма выбора моделей для синтеза и верификации режима 3.5

В данном режиме имеется много различных методов верификации моделей, но в данной работе целесообразно использовать параметры по



Фрагменты самих созданных статистических и системно-когнитивных моделей (СК-модели) приведены на рисунках (Рисунок 8, Рисунок 9, Рисунок 10).

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

Код признака	Наименование описательной шкалы и градации	1. OCCUPANCY NO	2. OCCUPANCY YES	Сумма	Среднее	Средн. квадр. откл.
1	DAYOFWEEK-Friday	1196	244	1440	720.00	673.17
2	DAYOFWEEK-Monday	1440		1440	720.00	1018.23
3	DAYOFWEEK-Saturday	946	495	1441	720.50	318.91
4	DAYOFWEEK-Sunday	1440		1440	720.00	1018.23
5	DAYOFWEEK-Thursday	1363	392	1755	877.50	686.60
6	DAYOFWEEK-Tuesday	1267	754	2021	1010.50	362.75
7	DAYOFWEEK-Wednesday	1744	1136	2880	1440.00	429.92
8	HOUR-1/10-{1.0000000, 3.2000000}	1620		1620	810.00	1145.51
9	HOUR-2/10-{3.2000000, 5.4000000}	1080		1080	540.00	763.68
10	HOUR-3/10-{5.4000000, 7.6000000}	1050	39	1089	544.50	714.88
11	HOUR-4/10-{7.6000000, 9.8000000}	501	530	1031	515.50	20.51
12	HOUR-5/10-{9.8000000, 12.0000000}	500	804	1304	652.00	214.96
13	HOUR-6/10-{12.0000000, 14.2000000}	452	440	892	446.00	8.49
14	HOUR-7/10-{14.2000000, 16.4000000}	315	774	1089	544.50	324.56
15	HOUR-8/10-{16.4000000, 18.6000000}	638	433	1071	535.50	144.96
16	HOUR-9/10-{18.6000000, 20.8000000}	1080		1080	540.00	763.68
17	HOUR-10/10-{20.8000000, 23.0000000}	1620		1620	810.00	1145.51
18	TEMPERATURE-1/10-{19.9908333, 19.9908333}	965		965	482.50	682.36
19	TEMPERATURE-2/10-{19.9908333, 20.4816667}	2659	18	2677	1338.50	1867.47
20	TEMPERATURE-3/10-{20.4816667, 20.9725000}	3443	204	3647	1823.50	2290.32
21	TEMPERATURE-4/10-{20.9725000, 21.4633333}	1133	378	1511	755.50	533.87
22	TEMPERATURE-5/10-{21.4633333, 21.9541667}	485	923	1408	704.00	309.71
23	TEMPERATURE-6/10-{21.9541667, 22.4450000}	189	510	699	349.50	226.98
24	TEMPERATURE-7/10-{22.4450000, 22.9358333}	166	365	531	265.50	140.71

Рисунок 8 – Матрица абсолютных частот (фрагмент)

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1"

Код признака	Наименование описательной шкалы и градации	1. OCCUPANCY NO	2. OCCUPANCY YES	Сумма	Среднее	Средн. квадр. откл.
1	DAYOFWEEK-Friday	0.011	-0.041	-0.030	-0.015	0.037
2	DAYOFWEEK-Monday	0.028		0.028	0.014	0.020
3	DAYOFWEEK-Saturday	-0.010	0.022	0.012	0.006	0.022
4	DAYOFWEEK-Sunday	0.028		0.028	0.014	0.020
5	DAYOFWEEK-Thursday	0.005	-0.016	-0.011	-0.006	0.015
6	DAYOFWEEK-Tuesday	-0.014	0.029	0.016	0.008	0.030
7	DAYOFWEEK-Wednesday	-0.017	0.034	0.017	0.009	0.036
8	HOUR-1/10-{1.0000000, 3.2000000}	0.028		0.028	0.014	0.020
9	HOUR-2/10-{3.2000000, 5.4000000}	0.028		0.028	0.014	0.020
10	HOUR-3/10-{5.4000000, 7.6000000}	0.024	-0.179	-0.154	-0.077	0.144
11	HOUR-4/10-{7.6000000, 9.8000000}	-0.036	0.058	0.021	0.011	0.066
12	HOUR-5/10-{9.8000000, 12.0000000}	-0.057	0.074	0.016	0.008	0.093
13	HOUR-6/10-{12.0000000, 14.2000000}	-0.033	0.054	0.021	0.011	0.061
14	HOUR-7/10-{14.2000000, 16.4000000}	-0.082	0.086	0.004	0.002	0.119
15	HOUR-8/10-{16.4000000, 18.6000000}	-0.018	0.036	0.018	0.009	0.039
16	HOUR-9/10-{18.6000000, 20.8000000}	0.028		0.028	0.014	0.020
17	HOUR-10/10-{20.8000000, 23.0000000}	0.028		0.028	0.014	0.020
18	TEMPERATURE-1/10-{19.9908333, 19.9908333}	0.028		0.028	0.014	0.020
19	TEMPERATURE-2/10-{19.9908333, 20.4816667}	0.027	-0.327	-0.300	-0.150	0.250
20	TEMPERATURE-3/10-{20.4816667, 20.9725000}	0.023	-0.139	-0.116	-0.058	0.114
21	TEMPERATURE-4/10-{20.9725000, 21.4633333}	0.002	-0.006	-0.004	-0.002	0.006
22	TEMPERATURE-5/10-{21.4633333, 21.9541667}	-0.067	0.079	0.012	0.006	0.103
23	TEMPERATURE-6/10-{21.9541667, 22.4450000}	-0.088	0.089	0.000	0.000	0.125
24	TEMPERATURE-7/10-{22.4450000, 22.9358333}	-0.075	0.083	0.008	0.004	0.112

Рисунок 9 – Матрица информативностей INF1 (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. OCCUPANCY NO	2. OCCUPANCY YES	Сумма	Среднее	Средн. квадр. откл.
1	DAYOFWEEK-Friday	142.573	-142.573			201.629
2	DAYOFWEEK-Monday	386.573	-386.573			546.697
3	DAYOFWEEK-Saturday	-108.159	108.159			152.959
4	DAYOFWEEK-Sunday	386.573	-386.573			546.697
5	DAYOFWEEK-Thursday	79.136	-79.136			111.915
6	DAYOFWEEK-Tuesday	-211.456	211.456			299.044
7	DAYOFWEEK-Wednesday	-362.854	362.854			513.153
8	HOUR-1/10-{1.000000, 3.200000}	434.894	-434.894			615.034
9	HOUR-2/10-{3.200000, 5.400000}	289.930	-289.930			410.022
10	HOUR-3/10-{5.400000, 7.600000}	253.346	-253.346			358.285
11	HOUR-4/10-{7.600000, 9.800000}	-253.225	253.225			358.114
12	HOUR-5/10-{9.800000, 12.000000}	-453.937	453.937			641.964
13	HOUR-6/10-{12.000000, 14.200000}	-200.540	200.540			283.606
14	HOUR-7/10-{14.200000, 16.400000}	-481.654	481.654			681.162
15	HOUR-8/10-{16.400000, 18.600000}	-145.486	145.486			205.749
16	HOUR-9/10-{18.600000, 20.800000}	289.930	-289.930			410.022
17	HOUR-10/10-{20.800000, 23.000000}	434.894	-434.894			615.034
18	TEMPERATURE-1/10-{19.500000, 19.9908333}	259.058	-259.058			366.363
19	TEMPERATURE-2/10-{19.9908333, 20.4816667}	700.650	-700.650			990.868
20	TEMPERATURE-3/10-{20.4816667, 20.9725000}	775.050	-775.050			1096.086
21	TEMPERATURE-4/10-{20.9725000, 21.4633333}	27.633	-27.633			39.079
22	TEMPERATURE-5/10-{21.4633333, 21.9541667}	-545.018	545.018			770.771
23	TEMPERATURE-6/10-{21.9541667, 22.4450000}	-322.351	322.351			455.873
24	TEMPERATURE-7/10-{22.4450000, 22.9358333}	-222.451	222.451			314.594

Рисунок 10 – 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 [6]. Это и есть решение проблемы сопоставимости в АСК-анализе и системе «Эйдос», отличающее их от других интеллектуальных технологий.

## 1.4. 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 [7]. В режиме 3.4 системы «Эйдос» изучается достоверность каждой частной модели в соответствии с этими мерами достоверности.

Из рисунка (Рисунок 11) мы видим, что в данном интеллектуальном приложении по F-критерию Ван Ризбергера наиболее достоверной является СК-модель INF1 с интегральным критерием «Семантический резонанс знаний» ( $F=0,959$  при максимуме 1), что является довольно хорошим результатом, по критерию L1 проф. Е.В. Луценко наиболее достоверной также является СК-модель INF3 с интегральным критерием «Сумма знаний» ( $L1=0,975$  при максимуме 1), что также является хорошим результатом.

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма модул. истинно-полож. решений (S1P)	Сумма модул. истинно-отриц. решений (S1N)	Сумма модул. ложно-полож. решений (S2P)	Сумма модул. ложно-отриц. решений (S2N)
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс частот с обр...	12417	12143	7391	5026	274	0.707	0.978	0.821	6507.414	809.570	754	3430
1. ABS - частный критерий: количество встреч признаков "клас...	Сумма абс частот по признак...	12417	6707	152	12265	5710	0.354	0.540	0.427	4589.502			3430
2. PRC1 - частный критерий: усл. вероятность n-го признака сред...	Корреляция усл.отн частот с о...	12417	12143	7391	5026	274	0.707	0.978	0.821	6507.414	809.570	754	3472
2. PRC1 - частный критерий: усл. вероятность n-го признака сред...	Сумма усл.отн частот по приз...	12417	12417		12417		0.500	1.000	0.667	7393.851			3472
3. PRC2 - частный критерий: условная вероятность n-го признака...	Корреляция усл.отн частот с о...	12417	12143	7391	5026	274	0.707	0.978	0.821	6507.212	809.561	754	3356
3. PRC2 - частный критерий: условная вероятность n-го признака...	Сумма усл.отн частот по приз...	12417	12417		12417		0.500	1.000	0.667	6750.659			3356
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	12417	11876	11934	483	541	0.961	0.956	0.959	2724.462	3729.397	59	63
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	12417	10493	11808	609	1924	0.945	0.845	0.892	1134.855	3504.463	63	62
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	12417	11862	11930	487	555	0.961	0.955	0.958	2746.872	3771.927	54	100
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	12417	10201	11397	1020	2216	0.909	0.822	0.863	1221.914	3425.041	100	199
6. INF3 - частный критерий: Унквадат, разности между фактич...	Семантический резонанс зна...	12417	11356	11356	1061	1061	0.915	0.915	0.915	5784.051	5784.051	199	53
6. INF3 - частный критерий: Унквадат, разности между фактич...	Сумма знаний	12417	11356	11356	1061	1061	0.915	0.915	0.915	5529.983	372.883	199	327
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	12417	11337	12121	296	1080	0.975	0.913	0.943	4724.212	6466.018	53	327
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	12417	11046	10551	1866	1371	0.855	0.890	0.872	2706.717	1937.323	54	186
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	12417	11320	12107	310	1097	0.973	0.912	0.941	4676.107	6437.519	54	186
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	12417	10625	10024	2393	1792	0.816	0.856	0.835	2495.883	1478.745	416	186
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Семантический резонанс зна...	12417	11306	11440	977	1111	0.920	0.911	0.915	5678.882	5675.039	231	192
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Сумма знаний	12417	11356	10991	1436	1061	0.898	0.915	0.901	3131.531	3567.145	231	192
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Семантический резонанс зна...	12417	11268	11395	1022	1149	0.917	0.907	0.912	5272.345	5257.532	192	295
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Сумма знаний	12417	11145	10724	1693	1272	0.868	0.898	0.883	2721.040	2731.162	295	295

Рисунок 11 – Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергера и L1- и L2-критериям проф. Е.В. Луценко (фрагмент)

Хотелось бы обратить внимание, что статистические модели, как правило, дают более низкую средневзвешенную достоверность

идентификации и не идентификации, чем модели знаний, и практически никогда – более высокую. Этим и оправдано применение моделей знаний и интеллектуальных технологий.

Это позволяет сделать обоснованные выводы о том, что:

– существует достаточно сильная причинно-следственная зависимость между данными датчиков и присутствием человека в помещении;

– эта причинно-следственная зависимость адекватно отражена в системно-когнитивных моделях INF1 и INF3;

– так как L2-мера проф. Е.В. Луценко для модели INF1 составляет 0,620, а для INF3 – 0,787, то предпочтение отдаем модели INF3.

На рисунке (Рисунок 12) приведено частотное распределение числа истинных и ложных положительных и отрицательных решений по результатам оценки данных датчиков в СК-модели INF3 по данным обучающей выборки.



Рисунок 12 – Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели INF3

На рисунке (Рисунок 13) изображены частотные распределения разностей долей разностей уровней сходимости решений в СК-модели INF3.

Исходя из данных на рисунках (Рисунок 12, Рисунок 13) можно сделать следующие выводы:

– левое и правое распределения практически равны по амплитуде. Левое распределение включает в себя как истинно-отрицательные, так и ложно-положительные решения, а правое, включает ложные отрицательные и истинно-положительные решения;

– сдвиг этих распределений относительно друг друга и другие различия между ними позволяют решать задачу присутствия человека в помещении на основании данных датчиков, а также другие задачи;



Рисунок 13 – Частотные распределения разностей долей разностей уровней сходимости решений в СК-модели INF3

– видно, что для отрицательных решений количество истинных решений всегда значительно превосходит количество ложных решений, причем при уровнях различия больше примерно 17% ложные отрицательные решения вообще практически отсутствуют;

– также видно, что для положительных решений при уровнях сходства от 0% до примерно 62% есть и истинные и ложные положительные решения, но число истинных решений больше числа ложных и доля истинных решений возрастает при увеличении уровня сходства. При уровнях сходства выше 62% ложно-положительных решений практически не встречается;

– модель INF3 немного лучше определяет непринадлежность объекта к классу, чем принадлежность (процент правильной не идентификации выше).

Любые данные о наблюдениях можно считать суммой истинного значения и шума, причем ни первое, ни второе неизвестны. Поэтому имеет смысл сравнить созданные модели с чисто случайными моделями, совпадающими по основным характеристикам. В системе «Эйдос» есть лабораторная работа № 2.01: «Исследование RND-модели при различных объемах выборки». Если данная работа устанавливается при отсутствии текущего приложения, то все параметры создаваемых моделей задаются вручную, если же текущая модель существует, как в нашем случае, то все основные ее параметры определяются автоматически (Рисунок 14).



Рисунок 14 – Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке (Рисунок 16) показано частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF3. Совершенно очевидное различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели (Рисунок 15) объясняется тем, что в реальных моделях кроме шума есть также и информация об истинных

причинно-следственных взаимосвязях факторов и их значений с одной стороны, и состояниями объекта моделирования, которые ими обуславливаются, с другой стороны. Если же такой информации в модели нет, то и распределение получается типа, приведенного на рисунке (Рисунок 16).

4.1.3.6. Обобщенная форма по достов.моделям при крит.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	Полнота модели	Фигера Ван Ризбергера	Сумма модул. уровней сход. истинно-поло... решений (STP)	Сумма модул. уровней сход. истинно-отриц. решений (STN)	Сумма модул. уровней сход. ложно-поло... решений
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс. частот с обр...	12417	6681	6206	6211	5736	0.518	0.538	0.528	1336.329	1157.040	1186
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс. частот по признак...	12417	12417		12417		0.500	1.000	0.667	11046.937		11030
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Корреляция усл.отн. частот с о...	12417	6682	6204	6213	5735	0.518	0.538	0.528	1336.329	1157.040	1186
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Сумма усл.отн. частот по при...	12417	12417		12417		0.500	1.000	0.667	11045.991		11029
3. PRC2 - частный критерий: условная вероятность иго признака...	Корреляция усл.отн. частот с о...	12417	6683	6205	6212	5734	0.518	0.538	0.528	1336.395	1157.098	1186
3. PRC2 - частный критерий: условная вероятность иго признака...	Сумма усл.отн. частот по при...	12417	12417		12417		0.500	1.000	0.667	11037.157		11020
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	12417	6635	6593	5824	5782	0.533	0.534	0.533	1466.259	1467.513	1231
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	12417	6452	6769	5648	5965	0.533	0.520	0.526	1197.840	1306.474	1002
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	12417	6635	6593	5824	5782	0.533	0.534	0.533	1466.287	1467.540	1231
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	12417	6422	6731	5686	5995	0.530	0.517	0.524	1178.522	1284.663	986
6. INF3 - частный критерий: X <sup>2</sup> и квадрат, разности между факти...	Семантический резонанс зна...	12417	6609	6609	5808	5808	0.532	0.532	0.532	1476.065	1476.065	1238
6. INF3 - частный критерий: X <sup>2</sup> и квадрат, разности между факти...	Сумма знаний	12417	6609	6609	5808	5808	0.532	0.532	0.532	1289.671	1289.671	1080
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	12417	6610	6610	5807	5807	0.532	0.532	0.532	1485.384	1485.384	1246
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	12417	6607	6607	5810	5810	0.532	0.532	0.532	1293.286	1293.286	1084
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	12417	6610	6610	5807	5807	0.532	0.532	0.532	1485.386	1485.386	1246
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	12417	6580	6580	5837	5837	0.530	0.530	0.530	1270.393	1270.187	1065
9. INF5 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Семантический резонанс зна...	12417	6609	6609	5808	5808	0.532	0.532	0.532	1476.064	1476.064	1238
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Сумма знаний	12417	6609	6609	5808	5808	0.532	0.532	0.532	1289.560	1289.561	1080
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Семантический резонанс зна...	12417	6609	6609	5808	5808	0.532	0.532	0.532	1475.851	1475.851	1238
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Сумма знаний	12417	6587	6587	5830	5830	0.530	0.530	0.530	1266.405	1266.209	1061

Помощь по меркам достоверности | Помощь по частотным распределениям | TP,TN,FP, FN | (TP-FP)/(TN-FN) | (T-F)/(T+F)\*100 | Задать интервал сглаживания

Рисунок 15 – Экранная форма с информацией о достоверности RND-моделей по F-критерию Ван Ризбергера и L1- и L2-критериям проф. Е.В. Луценко (фрагмент)

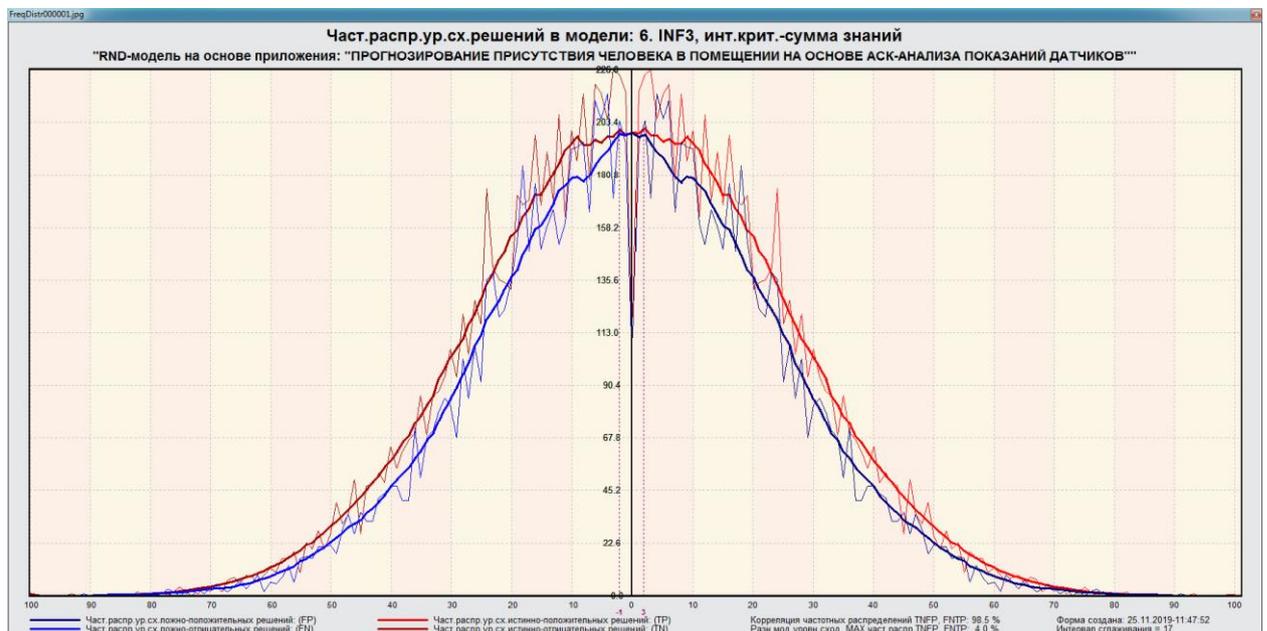


Рисунок 16 – Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в RND-модели INF3

На основе его сравнения с рисунком (Рисунок 12) можно сделать следующие выводы:

– RND-модель дает ложно-положительные и ложно-отрицательные прогнозы практически на всем диапазоне уровней сходства-различия;

– достоверность лучшей модели INF3, отражающей реальный объект моделирования, примерно на 50% выше, чем аналогичной случайной модели;

– различие между достоверностью статистических моделей и моделей знаний, созданных на основе случайной выборки, значительно меньше, чем у моделей, отражающих реальный объект моделирования;

– в реальных моделях кроме шума есть также и информация об истинных причинно-следственных взаимосвязях факторов и их значений с одной стороны, и состояниями объекта моделирования, которые ими обуславливаются, с другой стороны, причем примерно 1/3 достоверности обусловлена отражением в реальных моделях закономерностей предметной области, а 2/3 достоверности обусловлено наличием шума в исходных данных.

На основании этого можно предположить, что в исходных данных уровень сигнала о реальных причинно-следственных связях в моделируемой предметной области примерно в два раза ниже уровня шума [8].

## **2. Решение задач идентификации в наиболее достоверной модели**

### **2.1. Выбор наиболее достоверной модели и присвоение ей статуса текущей**

В соответствии со схемой обработки данных, информации и знаний в системе «Эйдос», присвоим СК-модели INF3 статус текущей модели. Для этого запустим режим 5.6 с параметрами, приведенными на экранной форме (Рисунок 17), результаты операции изображены на рисунке (Рисунок 18).

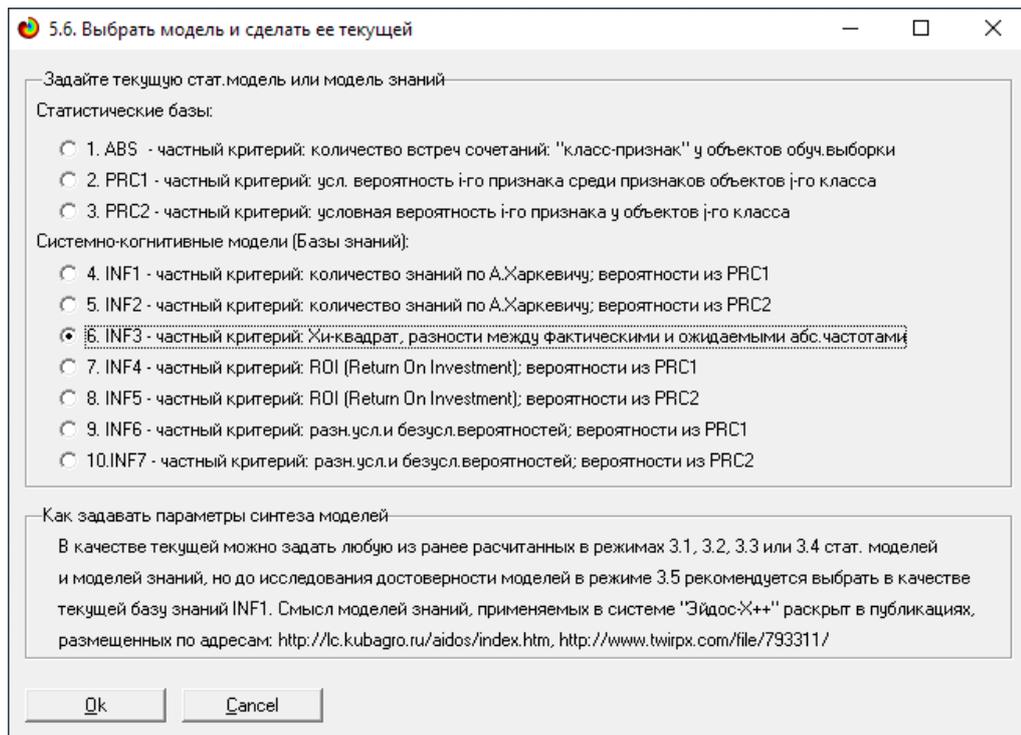


Рисунок 17 – Экранная форма придания наиболее достоверной по F-критерию СК-модели INF3 статуса текущей модели

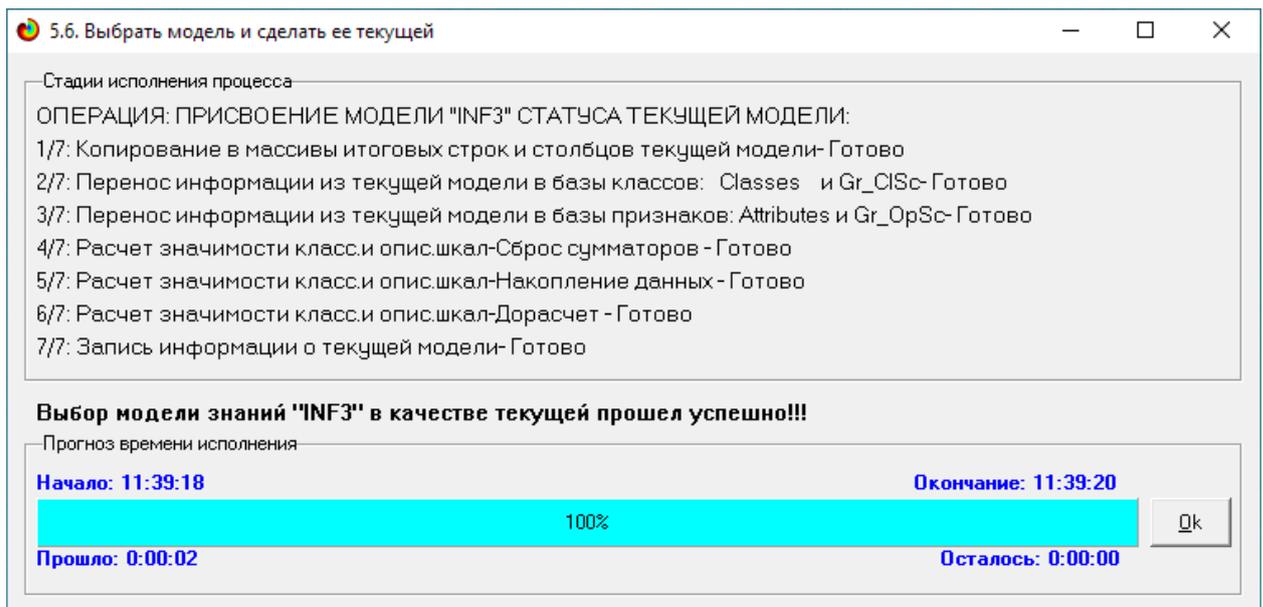


Рисунок 18 – Экранная формы результата присвоения статуса текущей модели

## 2.2. Прогнозирование (диагностика, классификация, распознавание, идентификация)

Решим задачу системной идентификации, т.е. определение присутствия человека на основе обучающей выборки в наиболее достоверной СК-модели INF3 (Рисунок 7).

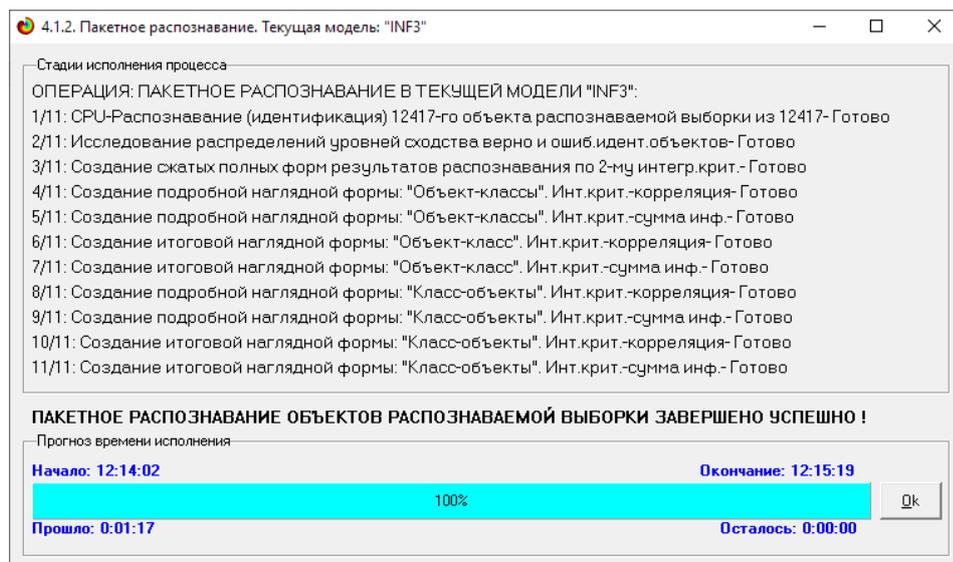


Рисунок 19 – Экранная форма с результатами работы пакетного распознавания модели INF3 (режим 4.1.2)

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;
- распознавание уровня сходства при разных моделях и интегральных критериях;
- достоверность идентификации классов при разных моделях и интегральных критериях.



Символ «v» стоит против тех результатов идентификации, которые подтвердились на опыте, т.е. соответствуют факту. Из рисунка (Рисунок 20) видно, что результаты идентификации являются хорошими, естественно при учете информации из рисунка (Рисунок 12) о том, что достоверные прогнозы в данной модели имеют уровень сходства выше 62%, т.е. по сути результаты с более низки уровнем сходства надо просто игнорировать.

### **2.3. Визуализация когнитивных функций**

Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора на переходы объекта управления в будущие состояния, соответствующие классам.

Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации 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; негативные, отражающие чему препятствуют система детерминации; 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) 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.

Применительно к задачам, рассматриваемым в данной работе, когнитивная функция показывает, какое количество информации содержится в различных значениях факторов о том, что объект моделирования перейдет в те или иные будущие состояния. В системе «Эйдос» за формирование когнитивных функций отвечает режим 4.5.

На рисунках (Рисунок 22, Рисунок 23, Рисунок 24, Рисунок 25, Рисунок 26, Рисунок 27, Рисунок 28) приведены примеры некоторых когнитивных функций, наглядно отражающих силу и направление влияния значений (т.е. степени выраженности) различных типов датчиков и других атрибутов на присутствие человека в помещении.

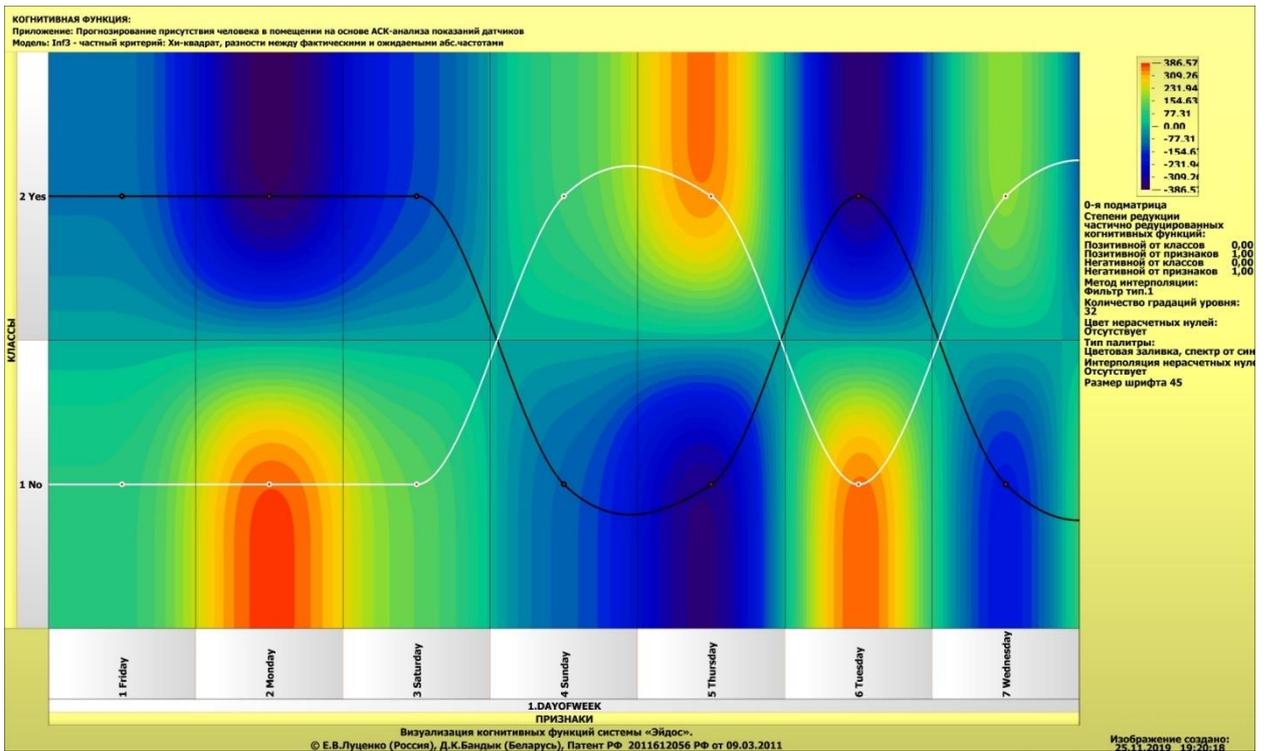


Рисунок 22 – Когнитивная функция, отражающая взаимосвязь присутствия человека в помещении и дня недели

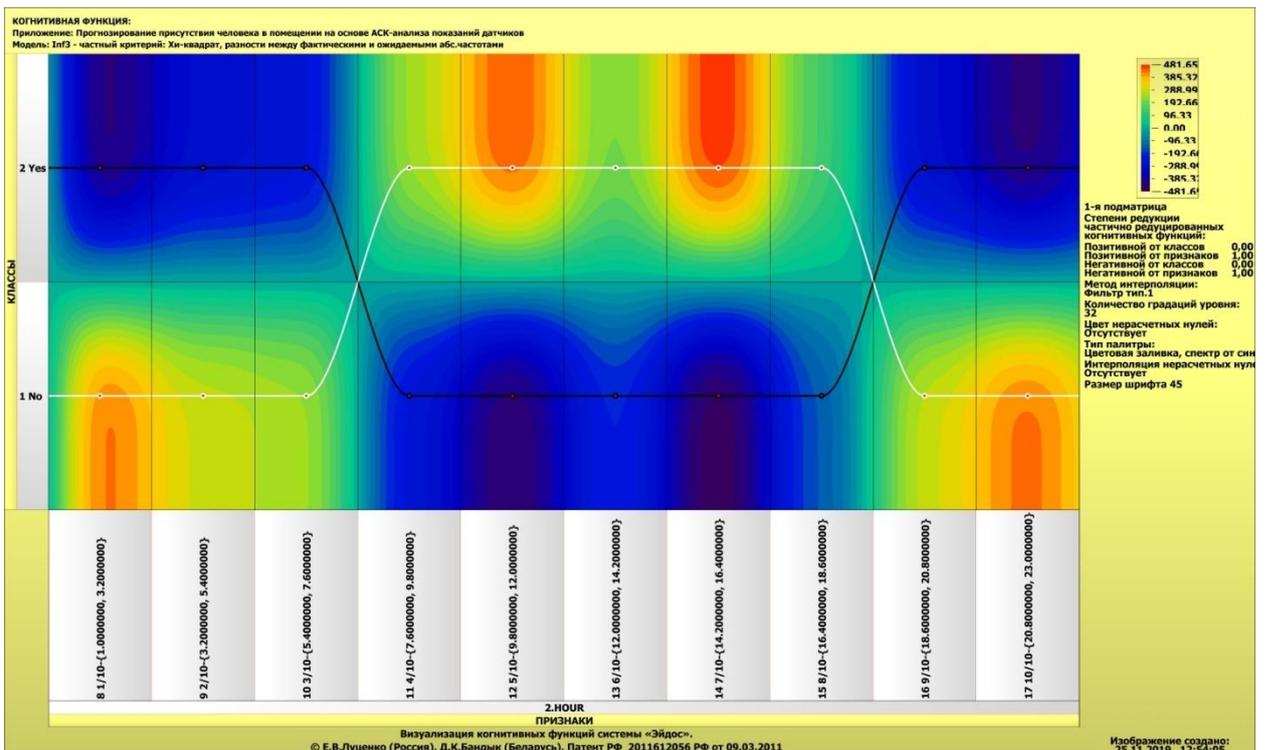


Рисунок 23 – Когнитивная функция, отражающая взаимосвязь присутствия человека в помещении и часа

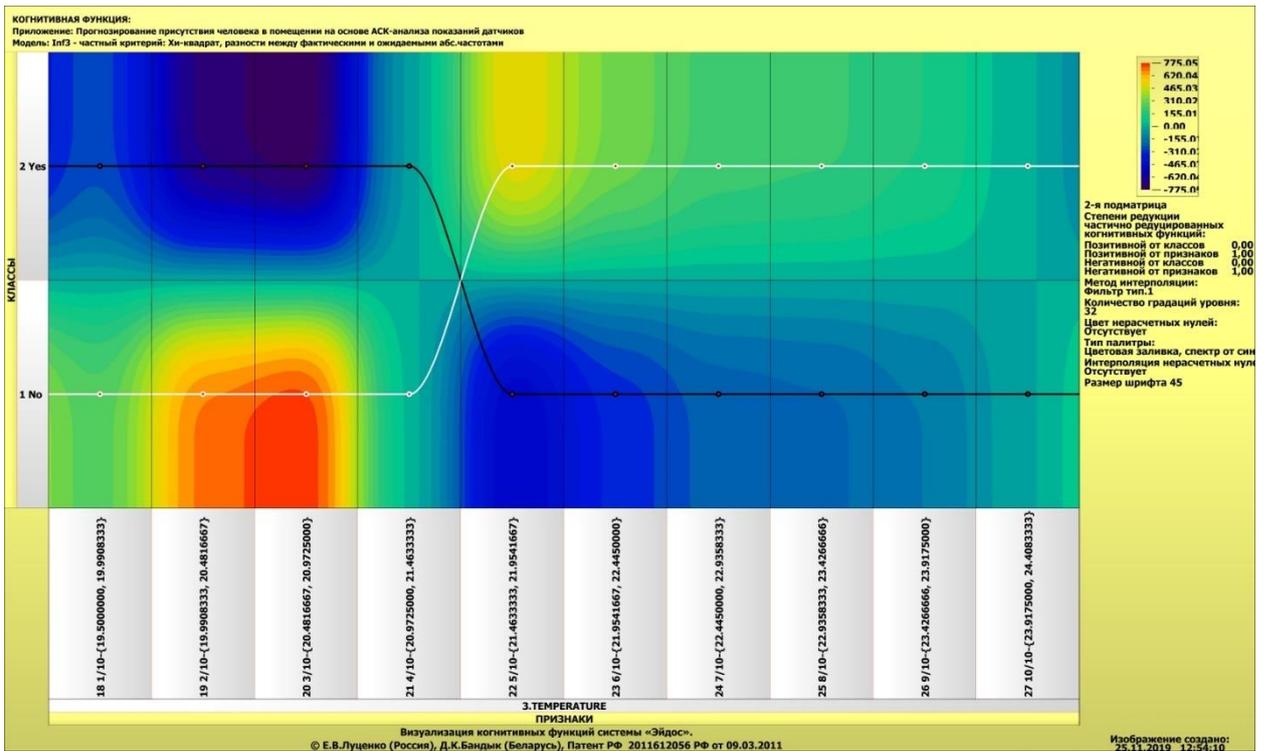


Рисунок 24 – Когнитивная функция, отражающая взаимосвязь присутствия человека в помещении и температуры

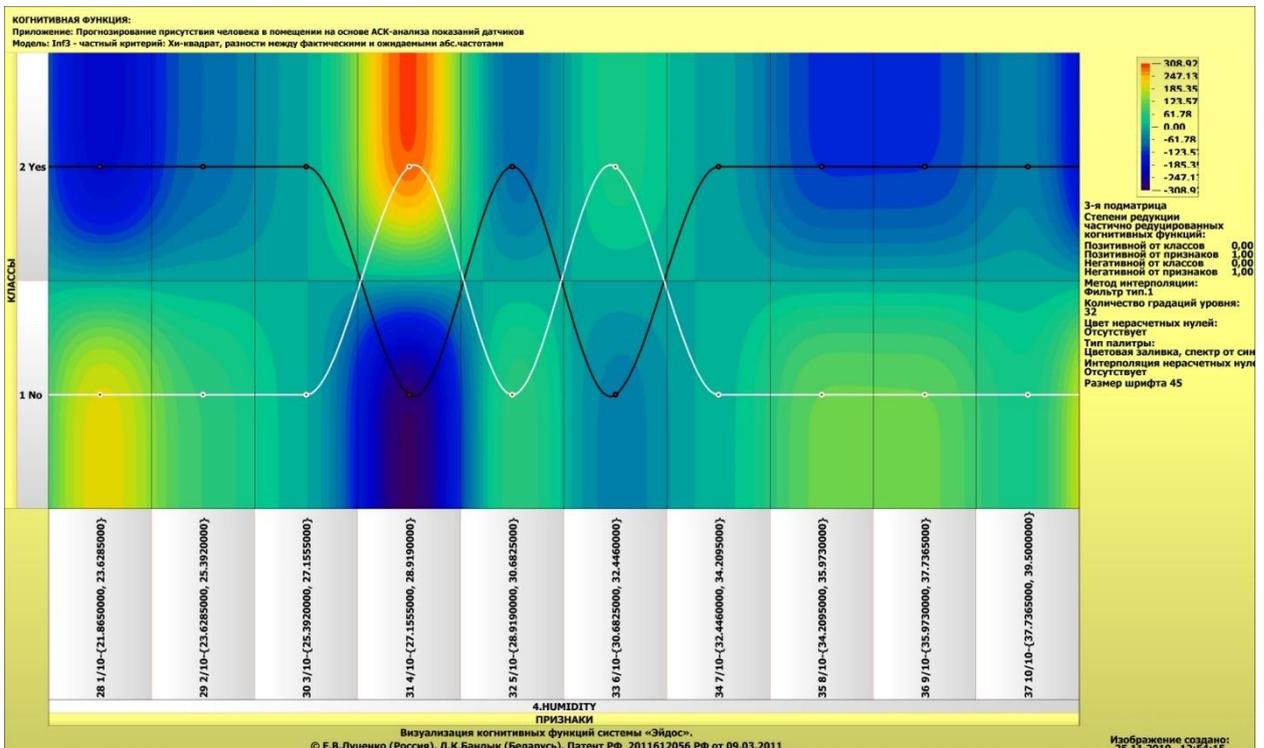


Рисунок 25 – Когнитивная функция, отражающая взаимосвязь присутствия человека в помещении и относительной влажности

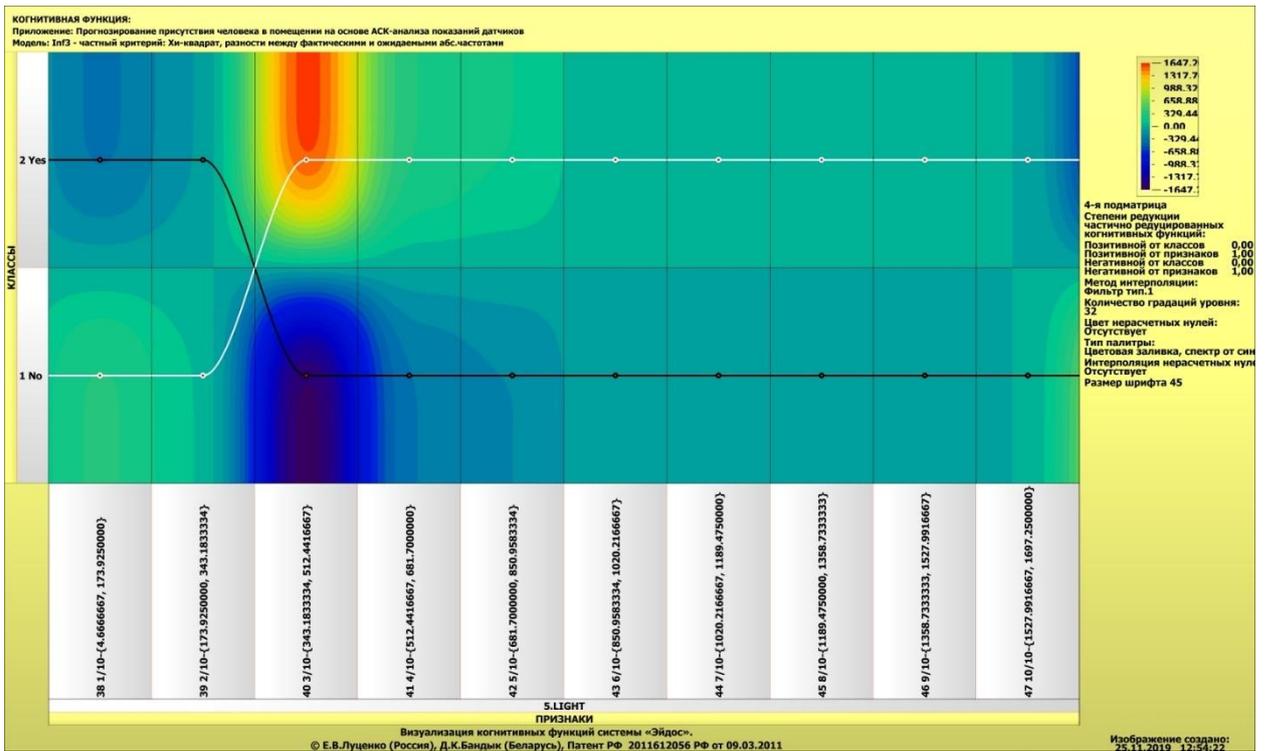


Рисунок 26 – Когнитивная функция, отражающая взаимосвязь присутствия человека в помещении и освещенности

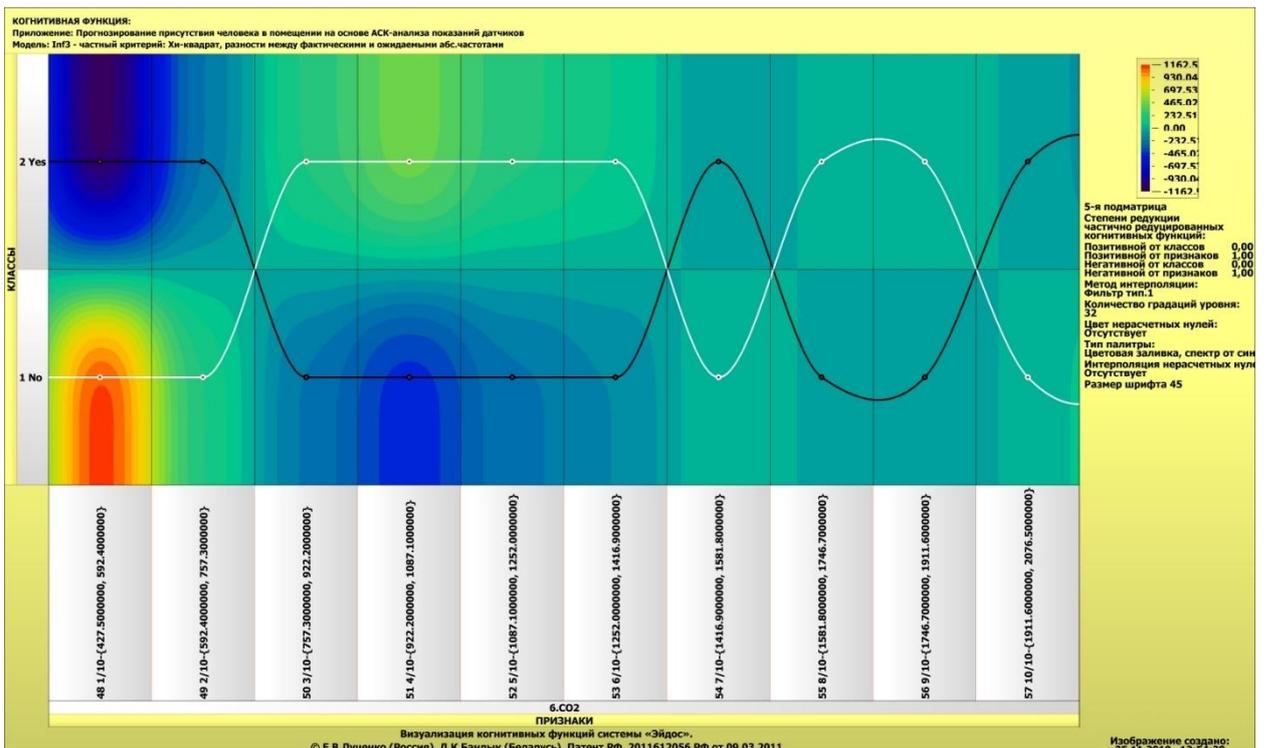


Рисунок 27 – Когнитивная функция, отражающая взаимосвязь присутствия человека в помещении и концентрации CO2

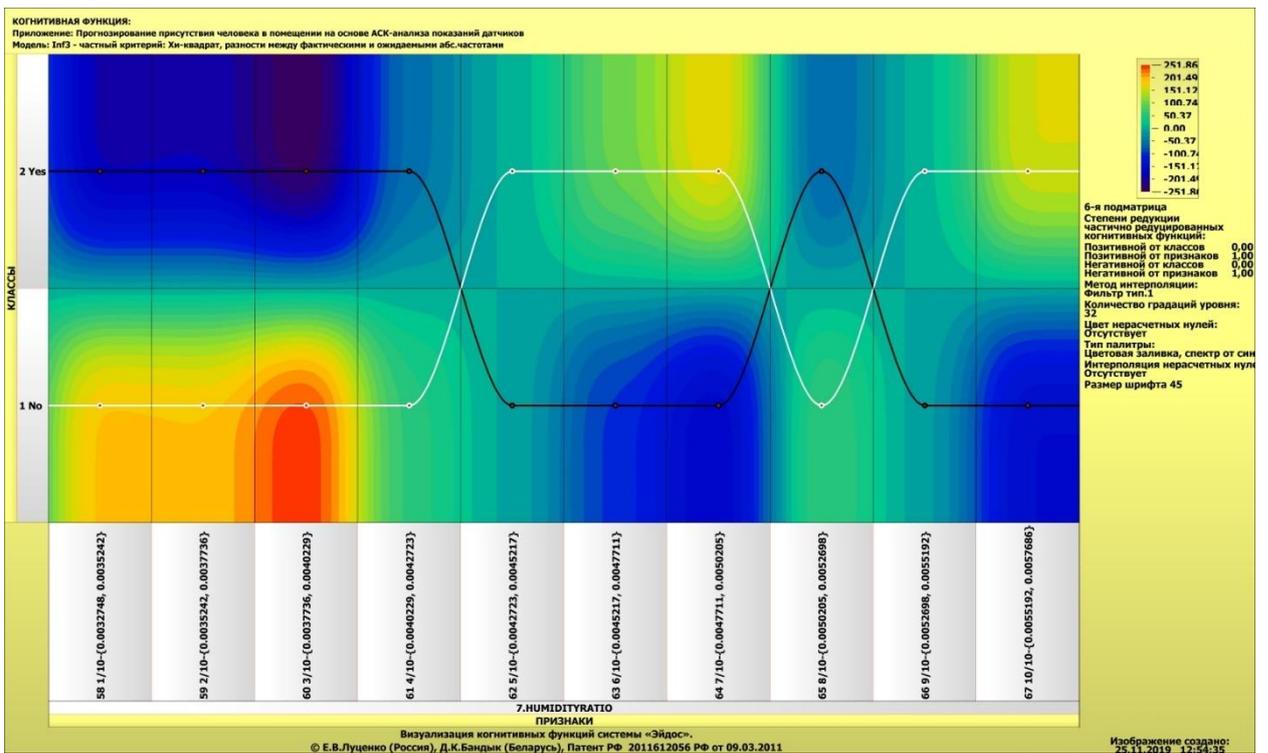


Рисунок 28 – Когнитивная функция, отражающая взаимосвязь присутствия человека в помещении и удельной влажности

В результате анализа полученных данных можно сделать следующие **ВЫВОДЫ**:

- когнитивная функция отражает зависимость дней недели от присутствия человека в помещении. В четверг хорошо прогнозируется присутствие, в понедельник и вторник – отсутствие (Рисунок 22);

- видно, что лучше всего принимаются решения о присутствии в рабочее время с 8 утра до 13 часов дня и с 14 часов дня до 17 часов вечера, а отсутствие лучше всего прогнозируется в ночное время 20 часов вечера до 4 часов утра, что соответствует истине, поскольку датасет был получен для офисного помещения (Рисунок 23);

- температура позволяет принять решение об отсутствии при значениях от 19 до 21 градуса, если температура выше, то ее показания нужно игнорировать (Рисунок 24);

- влажность имеет более сложную картину: лучше всего принимается решения о присутствии при показаниях 27-29%, остальные всплески скорее всего не дадут однозначной зависимости (Рисунок 25);

– освещенность позволяет принимать решения о присутствии в диапазоне 343-512 Люкс (Рисунок 26);

– концентрация углекислого газа позволяет принимать решения об отсутствии в диапазоне 427-592 ppm (Рисунок 27), поскольку организм человека выделяет CO<sub>2</sub> в процессе жизнедеятельности, а согласно санитарно-гигиеническими рекомендациями чистый городской воздух находится в диапазоне 400-600 ppm, поэтому показания выше 600 ppm могут говорить о присутствии человека в помещении (Рисунок 29);

– удельная влажность, по сравнению с относительной влажностью, дает более широкий диапазон предсказаний об отсутствии, уверенно предсказать отсутствие человека в помещении можно в диапазоне 0,00327 – 0,0040 (Рисунок 28).

### Carbon Dioxide (CO<sub>2</sub>) Hazard Scale

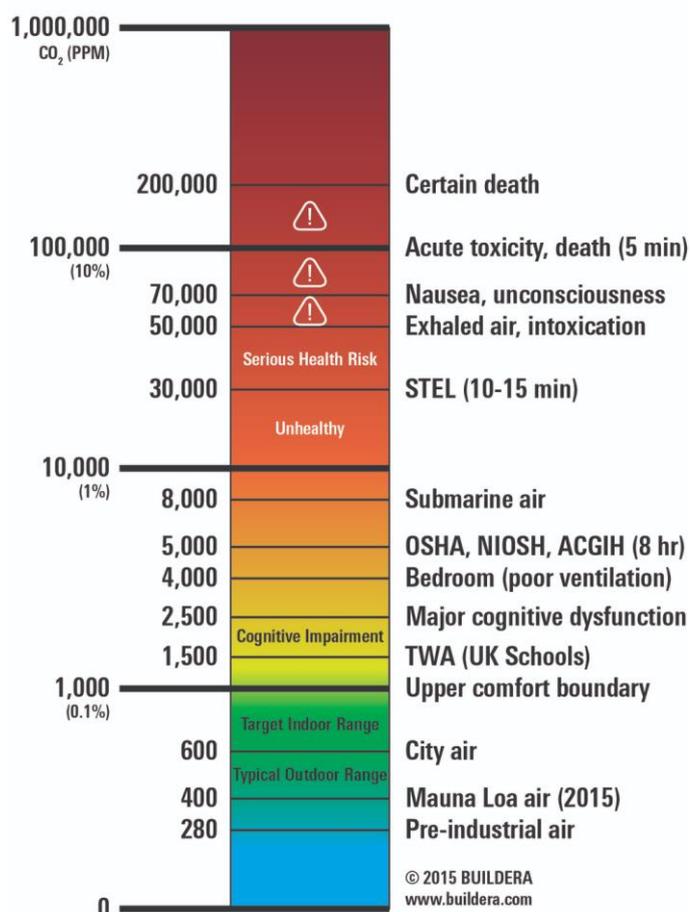


Рисунок 29 – Шкала опасности концентрации CO<sub>2</sub>

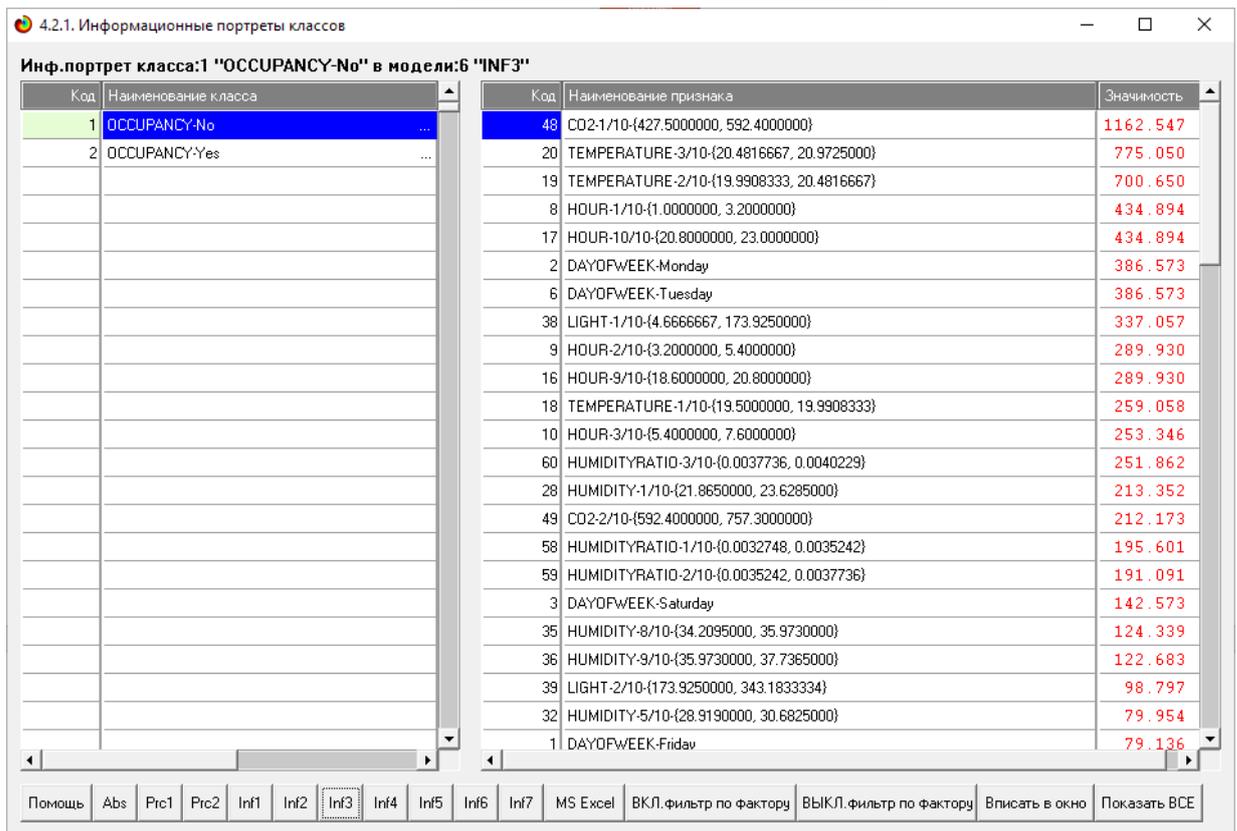


Рисунок 30 – Информационный портрет присутствия человека

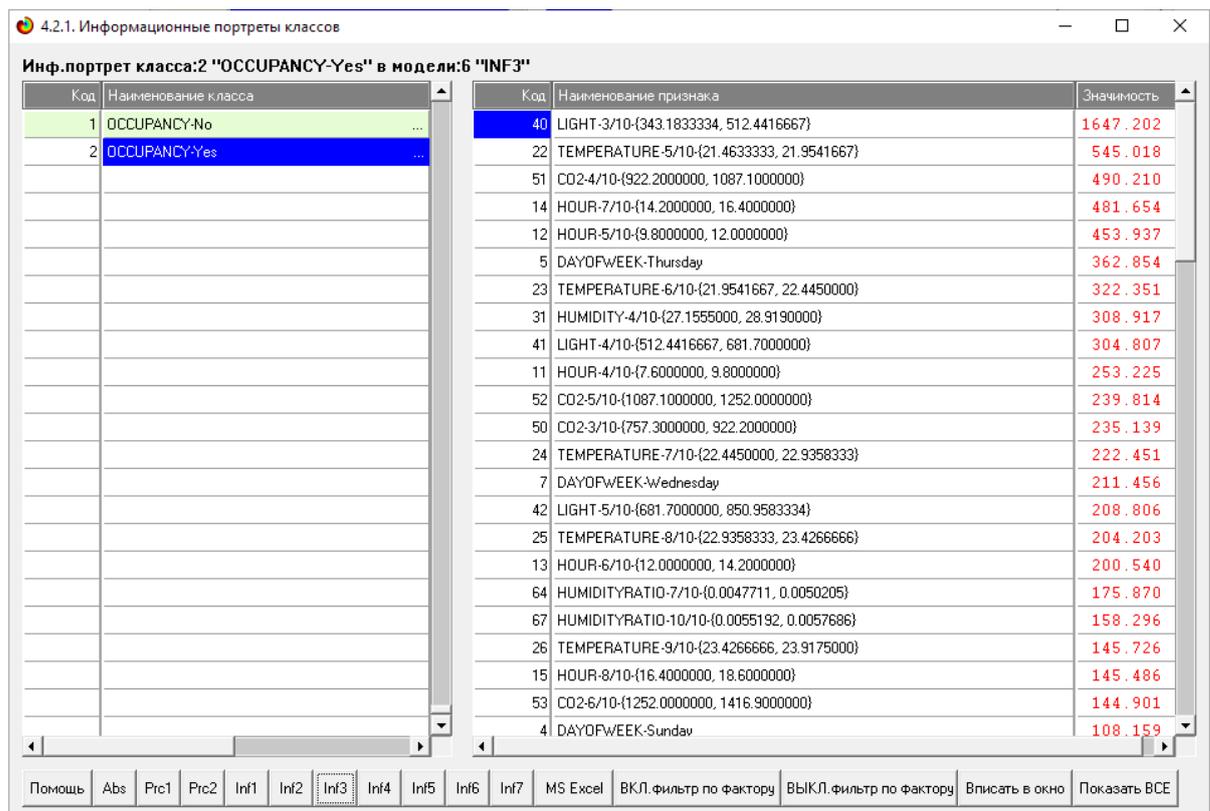


Рисунок 31 – Информационный портрет отсутствия человека

Согласно информационным портретам класса OCCUPANCY (Рисунок 30, Рисунок 31) на определения присутствия больше всего влияют показания

датчика концентрации углекислого газа (427-592 ppm), на отсутствие – датчик освещенности (343-512 Люкс).

## 2.4. Нелокальные нейроны и нейронная сеть

Для каждого технологического фактора в соответствии с предложенной моделью определяется величина и направление его влияния на осуществление всех желаемых и не желаемых хозяйственных ситуаций. Для каждой ситуации эта информация отображается в различных текстовых и графических формах, в частности, в форме нелокального нейрона (Рисунок 32, Рисунок 33). На данной диаграмме цвет линии означает знак связи (красный – положительная, синий – отрицательная), а толщина – ее модуль.

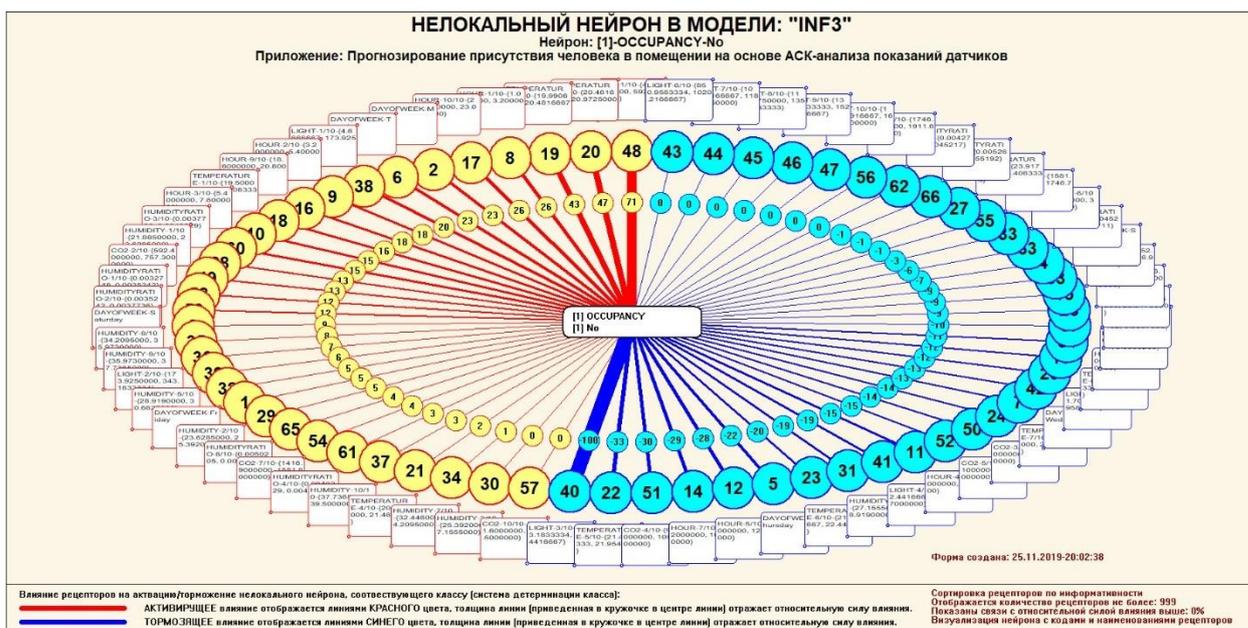


Рисунок 32 – Графическое отображение нелокальных нейронов в модели INF3 для нейрона OCCUPANCY-No

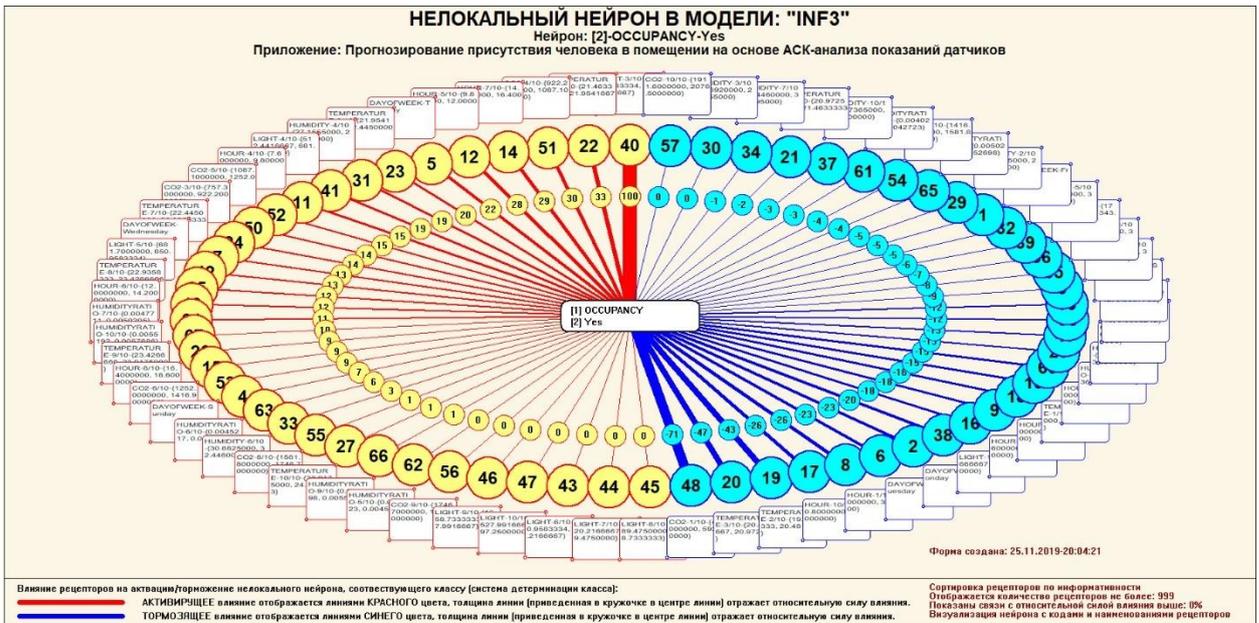


Рисунок 33 – Графическое отображение нелокальных нейронов в модели INF3 для нейрона OCCUPANCY-Yes

Дополнение модели нейрона связями факторов позволяет построить классическую когнитивную карту ситуации (будущего состояния АОУ). Детальная внутренняя структура любой связи отображается в форме инвертированной когнитивной диаграммы (Рисунок 34). Необходимо отметить, что все указанные графические формы генерируются системой «Эйдос» автоматически в соответствии с созданной моделью [10].

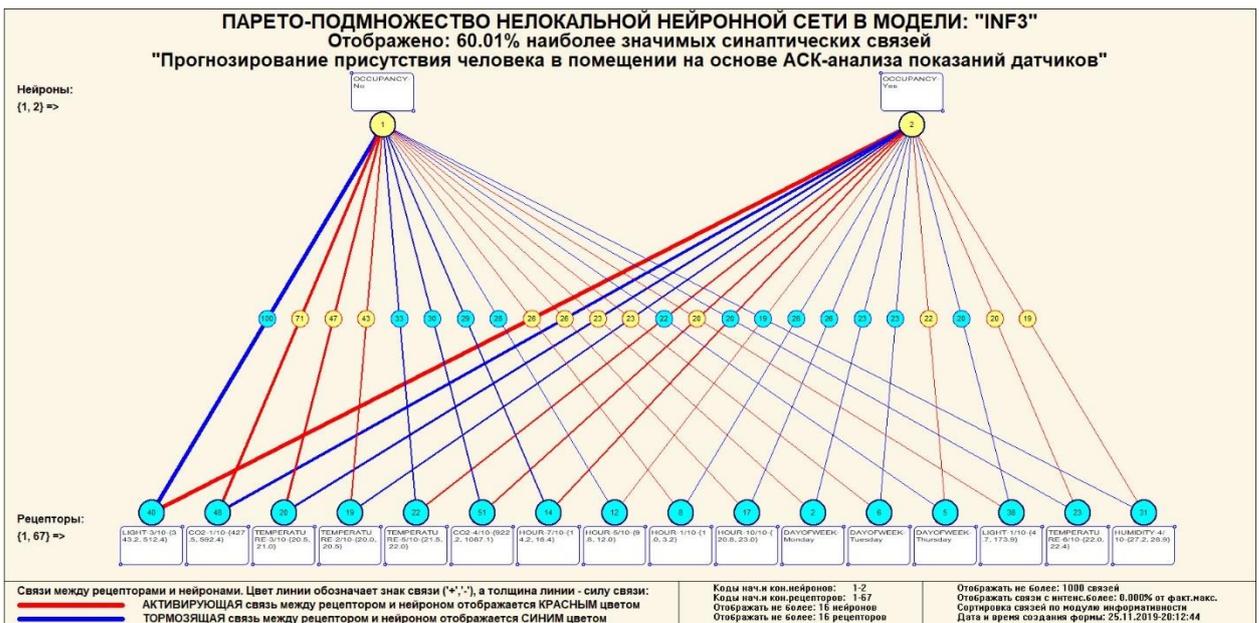


Рисунок 34 – Парето-подмножество нелокальной нейронной сети в модели INF3

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.

На рисунке (Рисунок 35) 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.

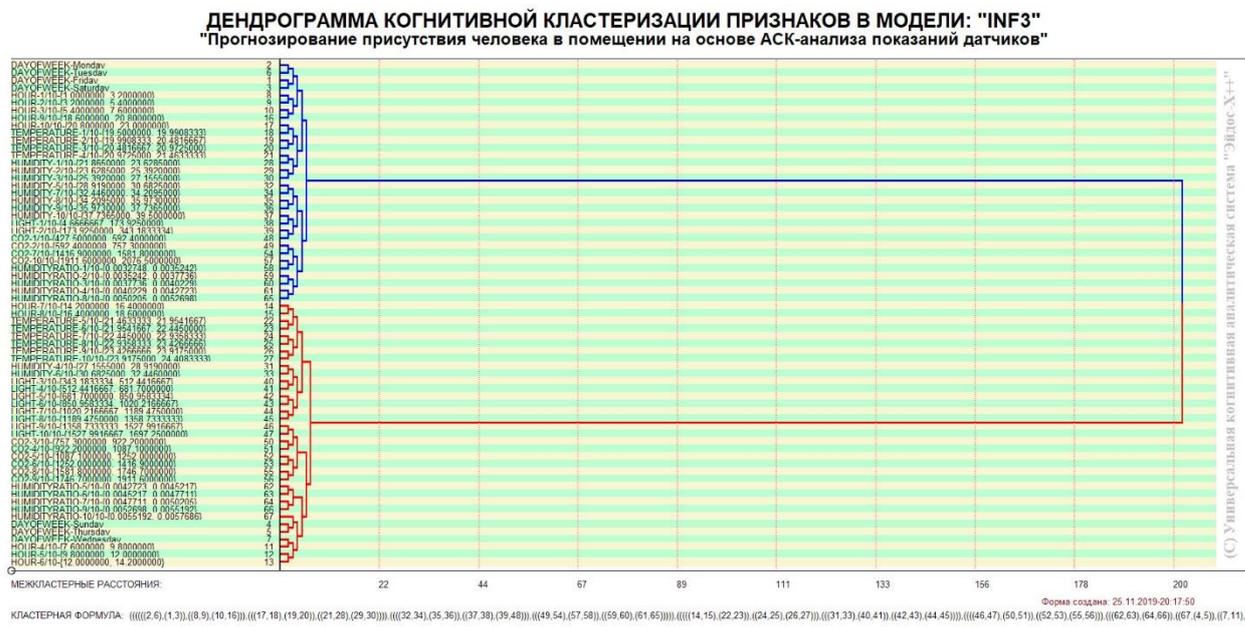


Рисунок 35 – Дендрограмма агломеративной когнитивной кластеризации значений параметров ДПН по их смыслу, т.е. по содержащейся в них информации о присутствии человека в помещении

## 2.5. 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 [11].

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают неформализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT-анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных.

Применительно к задаче, решаемой в данной работе, SWOT-анализ показывает степень влияния различных значений параметров на присутствие человека в помещении.

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 (Рисунок 36, Рисунок 37) приведены SWOT-диаграммы, наглядно отражающие силу и направление влияния различных значений параметров на присутствие человека в помещении.

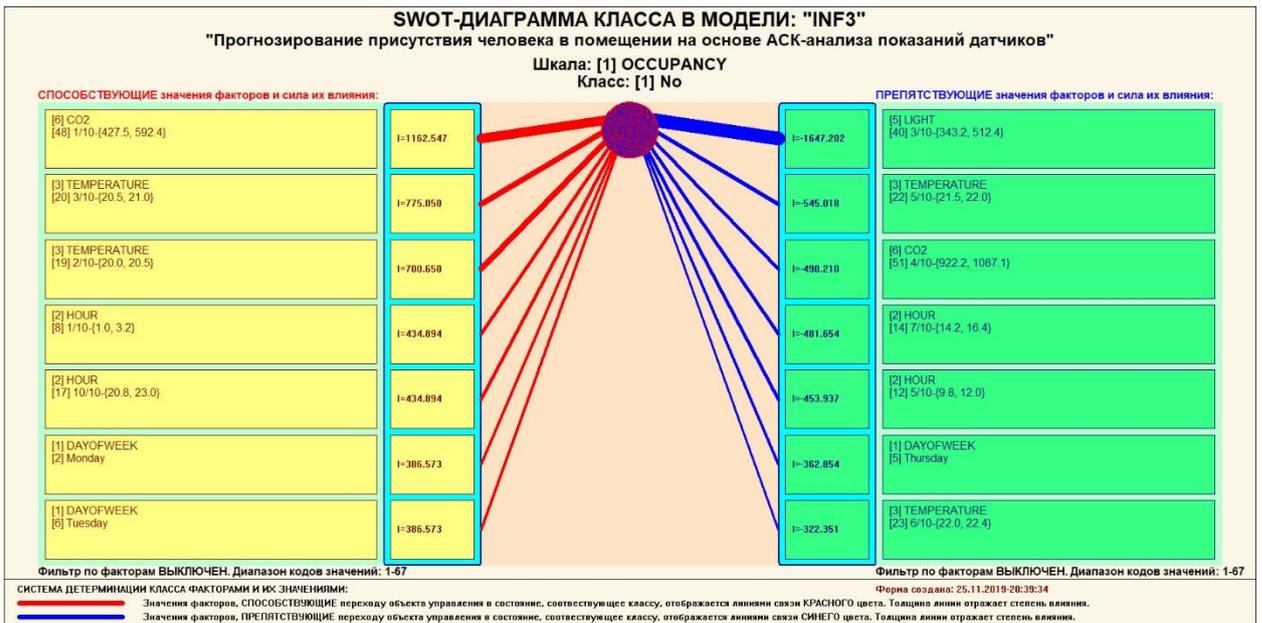


Рисунок 36 – SWOT-диаграмма класса OCCUPANCY-No в модели INF3

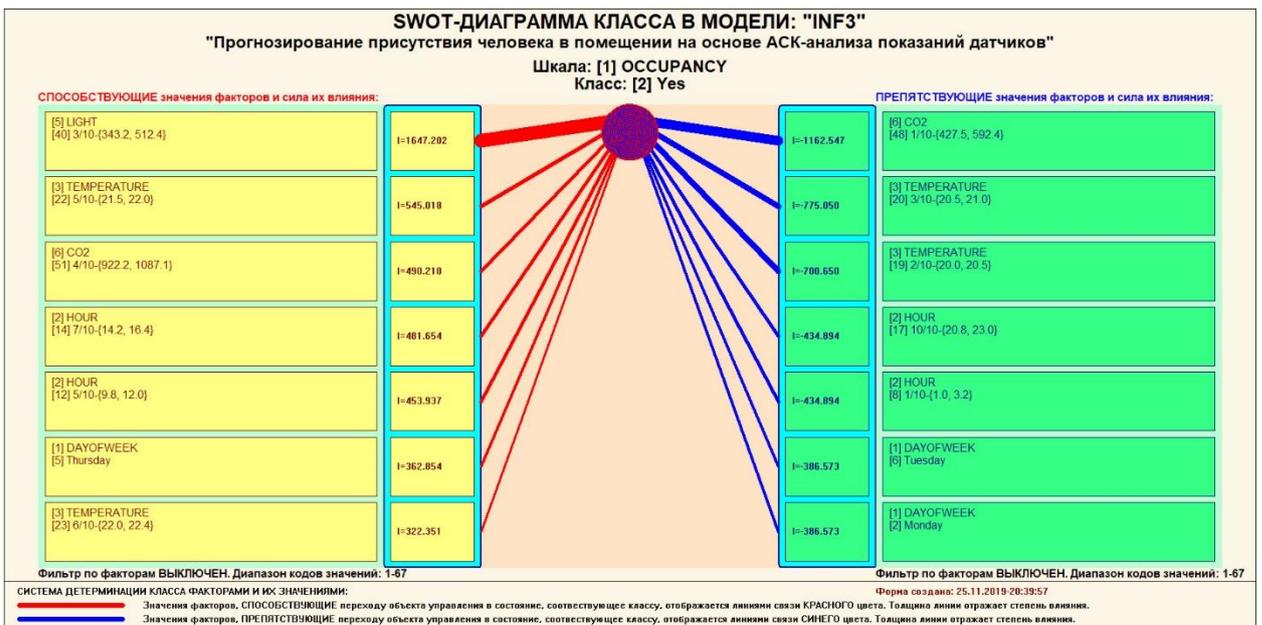


Рисунок 37 – SWOT-диаграмма класса OCCUPANCY-Yes в модели INF3

## Заключение

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.

Так как существует множество систем искусственного интеллекта, то возникает необходимость сопоставимой оценки качества их математических моделей. Одним из вариантов решения этой задачи является тестирование различных системы на общей базе исходных данных, для чего очень удобно использовать общедоступную базу репозитория UCI.

В данной курсовой работе приводится развернутый пример использования базы данных репозитория UCI для оценки качества математических моделей, применяемых в АСК-анализе и его программном инструментарии системе искусственного интеллекта «Эйдос». При этом наиболее достоверной в данном приложении оказались модель INF3, основанная на частном критерии Хи-квадрат, разности между фактическими и ожидаемыми абсолютными частотами при интегральном критерии «Сумма знаний». Точность модели составляет 0,915, что является отличным результатом. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием. Также необходимо обратить внимание, что статистические модели в данном приложении дают примерно на 21% более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, что, как правило, наблюдается и в других приложениях. Этим и оправдано применение моделей знаний.

В результате проделанной работы, с помощью системы «Эйдос» были созданы статистические и системно-когнитивные модели, в которых непосредственно на основе эмпирических данных сформированы обобщенные образы классов по различным значениям показателей присутствия человека в помещении, изучено влияние значений различных

параметров на эти прогнозы, и, на основе этого, решены задачи квалиметрии, классификации и исследования моделируемой предметной области путем исследования ее модели.

Данная работа является примером успешного применения когнитивных и информационных технологий для решения задач прогнозирования присутствия человека в помещении на основе АСК-анализа показаний датчиков.

## Список литературы

1. Occupancy Detection Data Set // [Электронный ресурс] intuit.ru. URL: <https://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+#>
2. Возможность использования систем автоматического управления освещением в длинном коридоре: XI межд. науч.-практ. конф. (Барнаул, 04-05.02.2016 г.) / Баранова И.А., Кондратьева Н.П., Большин Р.Г., Краснолуцкая М.Г., Коростелёв Д.В. Барнаул: АГАУ, 2016. С. 10-12.
3. Основы построения системы «Умный дом» // [Электронный ресурс] intuit.ru. URL: <https://www.intuit.ru/studies/courses/644/500/lecture/11366>
4. Сайт профессора Е.В. Луценко // [Электронный ресурс] intuit.ru. URL: <http://lc.kubagro.ru/>
5. Луценко Е.В., Боровко А.Ю. ПРОГНОЗИРОВАНИЕ КОЛИЧЕСТВА И КЛАССОВ СОЛНЕЧНЫХ ВСПЫШЕК НА ОСНОВЕ ИХ ПРЕДЫСТОРИИ ПО ДАННЫМ РЕПОЗИТОРИЯ UCI С ПРИМЕНЕНИЕМ АСК- АНАЛИЗА И ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ «ЭЙДОС» // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2014/10/pdf/99.pdf>
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. 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.

8. 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.

9. 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(67). С. 240 - 282. - Шифр Информрегистра: 0421100012\0077, 2,688 у.п.л. - Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.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. 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.

12. E. Hailemariam. Real-time occupancy detection using decision trees with multiple sensor types / Ebenezer Hailemariam, Rhys Goldstein, Ramtin Attar, Azam Khan // SimAUD '11 Proceedings of the 2011 Symposium on Simulation for Architecture and Urban Design, Pages 141-148. Boston, Massachusetts – April 03 - 07, 2011