

УДК 004.8

UDC 004.8

06.02.00 Ветеринария и Зоотехния

Veterinary and Zootechnics

**АГЛОМЕРАТИВНАЯ КОГНИТИВНАЯ
КЛАСТЕРИЗАЦИЯ СИМПТОМОВ И
СИНДРОМОВ В ВЕТЕРИНАРИИ****AGGLOMERATIVE COGNITIVE
CLUSTERING OF SYMPTOMS AND
SYNDROMES IN VETERINARY MEDICINE**

Луценко Евгений Вениаминович
д.э.н., к.т.н., профессор
Scopus Author ID: 57191193316
РИНЦ SPIN-код: 9523-7101
prof.lutsenko@gmail.com <http://ic.kubagro.ru>
Кубанский государственный аграрный университет имени И.Т.Трубилина, Краснодар, Россия

Lutsenko Evgeniy Veniaminovich
Dr.Sci.Econ., Cand.Tech.Sci., professor
Scopus Author ID: 57191193316
RSCI SPIN-code: 9523-7101
prof.lutsenko@gmail.com <http://ic.kubagro.ru>
Kuban State Agrarian University, Krasnodar, Russia

В статье на небольшом численном примере рассматривается сходство и различие симптомов и синдромов по их диагностическому смыслу, т.е. по той информации, которую они содержат о принадлежности состояний животных к различным нозологическим образам. Эта задача решается для ветеринарии с применением нового метода агломеративной когнитивной кластеризации, реализованного в автоматизированном системно-когнитивном анализе (АСК-анализ). Этот метод кластеризации отличается от известных традиционных тем, что: а) в нем параметры обобщенного образа кластера вычисляются не как средние от исходных объектов (симптомов) или их центр тяжести, а определяются с помощью той же самой базовой когнитивной операции АСК-анализа, которая применяется и для формирования обобщенных образов классов на основе примеров объектов и которая действительно корректно обеспечивает обобщение; б) в качестве критерия сходства используется не евклидово расстояние или его варианты, а интегральный критерий неметрической природы: «суммарное количество информации», применение которого теоретически корректно и дает хорошие результаты в неортонормированных пространствах, которые как правило и встречаются на практике; в) кластерный анализ проводится не на основе исходных переменных, матриц частот или матрицы сходства (различий), зависящих от единиц измерения по осям (измерительным шкалам), а в когнитивном пространстве, в котором по всем осям используется одна единица измерения: количество информации, и поэтому результаты кластеризации не зависят от исходных единиц измерения признаков объектов. Все это позволяет получить результаты кластеризации, понятные специалистам и поддающиеся содержательной интерпретации, хорошо согласующиеся с оценками экспертов, их опытом и интуитивными ожиданиями, что часто представляет собой проблему для классических методов кластеризации

In the article, on a small numerical example, we consider the similarity and difference of symptoms and syndromes according to their diagnostic meaning, i.e. according to the information they contain about the belonging of conditionals of animals to different nosological images. This problem can be solved for veterinary with the use of a new method of agglomerative cognitive clustering, implemented in Automated System-Cognitive analysis (ASC-analysis). This method of clustering differs from the known traditional methods in: a) in this method, the parameters of the generalized image of the cluster are calculated not as averages from the original objects (symptoms) or their center of gravity, but are determined using the same basic cognitive operation of ASC-analysis, which is used to form generalized images of the classes based on examples of objects and which really correctly provides a generalization; b) the similarity criterion is not the Euclidean distance or its variants, but the integral criterion of non-metric nature: "the total amount of information", the application of which is theoretically correct and gives good results in unortonormated spaces, which are usually found in practice; c) cluster analysis is carried out not on the basis of initial variables, frequency matrices or matrix of similarity (differences), depending on the units of measurement on the axes (measurement scales), but in cognitive space, in which one unit of measurement is used for all axes: the amount of information, and therefore the results of clustering do not depend on the initial units of measurement of features of objects. All this allows us to get the results of clustering, understandable to specialists and amenable to meaningful interpretation, well-consistent with the experts' assessments, their experience and intuitive expectations, which is often a problem for classical clustering methods

Ключевые слова: АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ,

Keywords: AUTOMATIVE SYSTEM-COGNITIVE ANALYSIS, INTELLECTUAL SYSTEM "EIDOS",

Doi: 10.21515/1990-4665-139-033

Задача нозологической диагностики в ветеринарии может решаться на основе знания *симптоматики*¹ и по сути представляет собой задачу классификации состояния объекта по его признакам (идентификация, распознавание).

Симптомы можно классифицировать на

– *детерминистские*, взаимно-однозначно связанные с определенными заболеваниями, т.е. встречающимися только при них и при них встречающихся всегда (эти симптомы наиболее ценны для диагностики);

– *статистические*, которые чаще наблюдаются при одних заболеваниях, а при других реже (эти симптомы менее ценны для диагностики, чем детерминистские);

– *бесполезные*, которые всегда встречаются при всех заболеваниях, но отсутствуют у здоровых животных (наличие этих симптомов позволяет установить лишь сам факт заболевания, но не позволяет установить какое оно именно).

Большинство симптомов относится к статистическим, а детерминистские и бесполезные симптомы представляют собой два полюса к которым в различной степени приближаются наиболее ценные и наименее ценные статистические симптомы.

Существование статистических симптомов приводит к ситуации, при которой один и тот же симптом несет информацию о различных заболеваниях, но само это количество информации разное, для разных заболеваний, т.е. об одних заболеваниях статистический симптом несет больше инфор-

¹ Симптом (греч. *symptōma* совпадение, случай, признак) признак болезни, обнаруживаемый с помощью клинических методов исследования и используемый для диагностики и (или) прогноза заболевания https://dic.academic.ru/dic.nsf/enc_medicine/28346/симптом

мации, а о других меньше. Таким образом, различные симптомы могут иметь сходный диагностический смысл, т.е. нести информацию о принадлежности состояния животного к одним и тем же *нозологическим² образам*. Группа симптомов, сходных по диагностическому смыслу образует *синдром³*.

В этой связи возникает **задача** количественного определения сходства-различия симптомов по их диагностическому смыслу, т.е. задача выявления синдромов.

Решение данной задачи представляет интерес как для ученых, разрабатывающих и совершенствующих методы и методики количественной диагностики в ветеринарии, так и для врачей-ветеринаров, применяющих эти методы диагностики на практике.

Для решения поставленной задачи может быть применен метод кластерного анализа.

Кластерный анализ (англ. *Data clustering*) – это задача разбиения заданной выборки *объектов* (ситуаций) на подмножества, называемые кластерами, так, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались.

Однако мы не будем применять классические методы кластерного анализа и реализующий их программный инструментарий. Причиной этого является то, что эти методы имеют ряд довольно существенных (на взгляд автора) недостатков. Эти недостатки подробно проанализированы в работе [1], в ней же предложено и их решение [1, 2]. Поэтому в данной статье мы не будем их подробно рассматривать. Отметим лишь, что многие из этих недостатков, являющихся проблемами классических методов кластерного анализа (таблица 1), преодолены в новом инновационном (т.е. доведенным до практического применения) методе искусственного интеллекта: автома-

² от греч. *posos* — болезнь и *logos* — слово, учение; букв.— учение о болезни.

³ Синдром (*syndromotom*, греч. *syndromē* стечение, скопление) устойчивая совокупность ряда симптомов с единым патогенезом: https://dic.academic.ru/dic.nsf/enc_medicine/28878/синдром

тизированном системно-когнитивном анализе (АСК-анализ) и его программном инструментарии: Универсальной автоматизированной системе «Эйдос» [3, 4]⁴.

Таблица 1 – ПРОБЛЕМЫ КЛАССИЧЕСКИХ МЕТОДОВ КЛАСТЕРИЗАЦИИ И ИХ РЕШЕНИЕ ПРЕДЛАГАЕМЫЕ В АСК-АНАЛИЗЕ И СИСТЕМЕ «ЭЙДОС»

№	Формулировка проблемы кластерного анализа	Предлагаемое в АСК-анализе и реализованное в системе «Эйдос» решение
1.	<i>Проблема 1.1</i> выбора метрики, корректной для неортонормированных пространств.	Предлагается применить неметрический интегральный критерий, представляющий собой суммарное количество информации в системе признаков о принадлежности объекта к классу («информационное расстояние»), никак не основанный на предположении об ортонормированности пространства и корректно работающий в неортонормированных пространствах
2.	<i>Проблема 1.2</i> ортонормирования пространства.	Поддерживается
3.	<i>Проблема 2.1</i> сопоставимой обработки описаний объектов, описанных признаками различной природы, измеряемыми в различных единицах измерения (проблема размерностей).	Объекты формально описываются в виде векторов, затем рассчитывается матрица абсолютных частот и на ее основе – матрица знаний, с использованием которой все признаки измеряются в одних единицах измерения: единицах измерения количества данных, информации и знаний – битах, байтах, и т.д.
	<i>Проблема 2.2</i> формализации описаний объектов, имеющих как количественные, так и качественные признаки.	Числовые шкалы преобразуются в интервальные значения, после чего градации всех типов шкал обрабатываются единообразно (см.п.3)
4.	<i>Проблема 3.1</i> доказательства гипотезы о нормальности исходных данных.	Нет необходимости, т.к. предлагаемые частные и интегральные критерии не предполагают нормальности исходных данных

⁴ http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

5.	<i>Проблема 3.2 нормализации исходных данных.</i>	Реализованы режимы ремонта или взвешивания исходных данных.
6.	<i>Проблема 3.3 применения непараметрических методов кластеризации, корректно работающих с ненормализованными данными.</i>	Предлагаемые методы являются непараметрическими и корректно работают с ненормализованными данными
7.	<i>Проблема 4 разработки такого метода кластерного анализа, математическая модель и алгоритм и которого органично включали бы фильтр, подавляющий шум в исходных данных, в результате чего данный метод кластеризации корректно работал бы при наличии шума в исходных данных.</i>	Предлагаемый метод включает фильтр подавления шума на уровне формирования матрицы абсолютных частот и самой математической форме интегрального критерия. Кроме того, реализованы режимы удаления или корректной обработки артефактов, выбросов (нетипичных объектов) и малопредставленных данных, по которым нет достаточной статистики в исходных данных
8.	<i>Проблема 5 разработки метода кластерного анализа, математическая модель и алгоритм и которого обеспечивали бы выявление «выбросов» (артефактов) в исходных данных и позволяли либо вообще не показывать их в дендрограммах, либо показывать, но так, чтобы было наглядно видно, что это артефакты.</i>	Поддерживается исключение выбросов и артефактов из дендрограмм, либо их отображение специальным для них образом.

В работе [1] были отмечены также и недостатки самого предложенного в ней решения, т.е. указаны перспективы его развития. Эти недостатки связаны прежде всего с программной реализацией DOS-версии самой системы «Эйдос», которая была текущей на момент написания работы [1]. В новой версии системы «Эйдос» от 05.06.2018 все эти ограничения DOS-версии преодолены и реализованы многие из указанных в [1] перспектив. В частности реализован режим агломеративной когнитивной кластеризации классов (режим 4.2.2.3) и признаков (режим 4.3.2.3).

Применим режим агломеративной когнитивной кластеризации признаков для решения поставленной в данной работе задачи. При этом мы используем интеллектуальное приложение системы «Эйдос», описанное в статьях [5, 6]. Отметим также, что данное приложение размещено в облаке на ftp-сервере системы «Эйдос» и является современным развитием ранее проведенной работы [7].

Вместо описания режима 4.3.2.2 приведем его Help (рисунок 1):

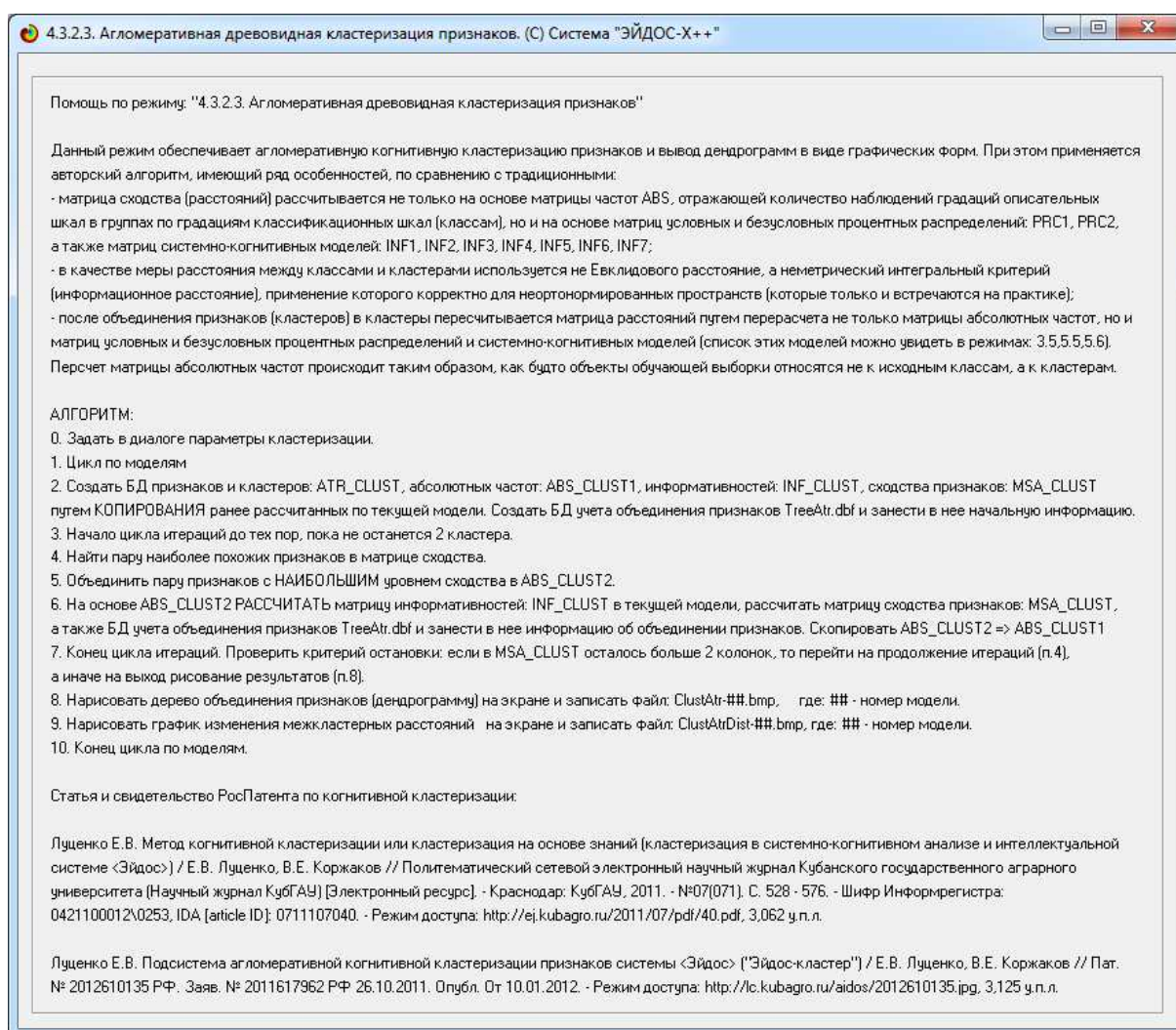
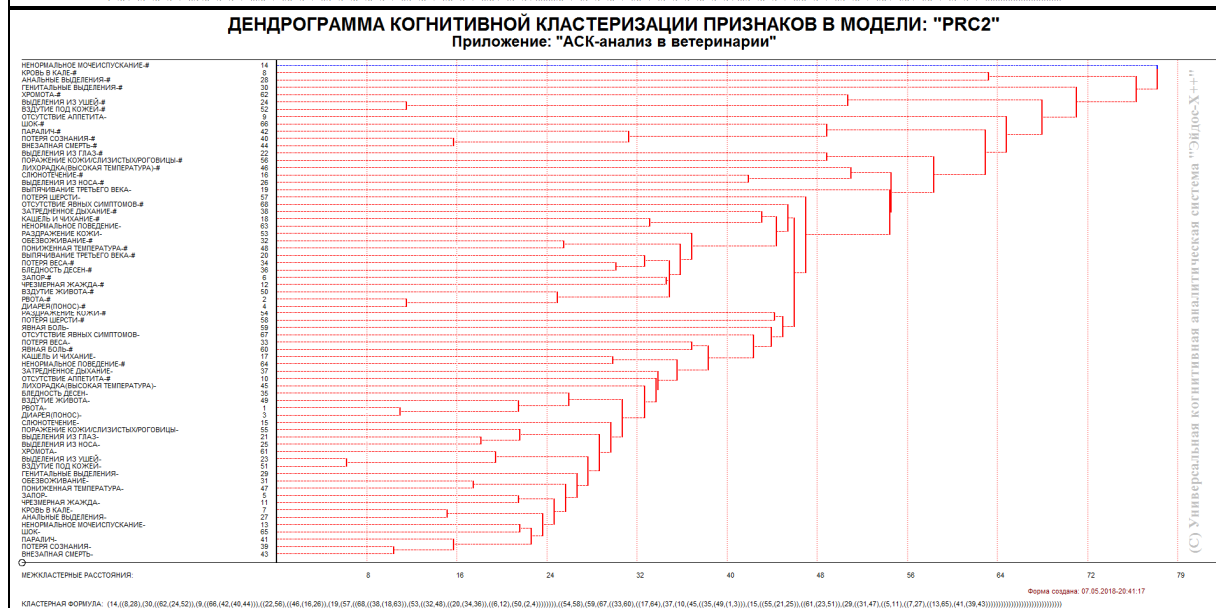
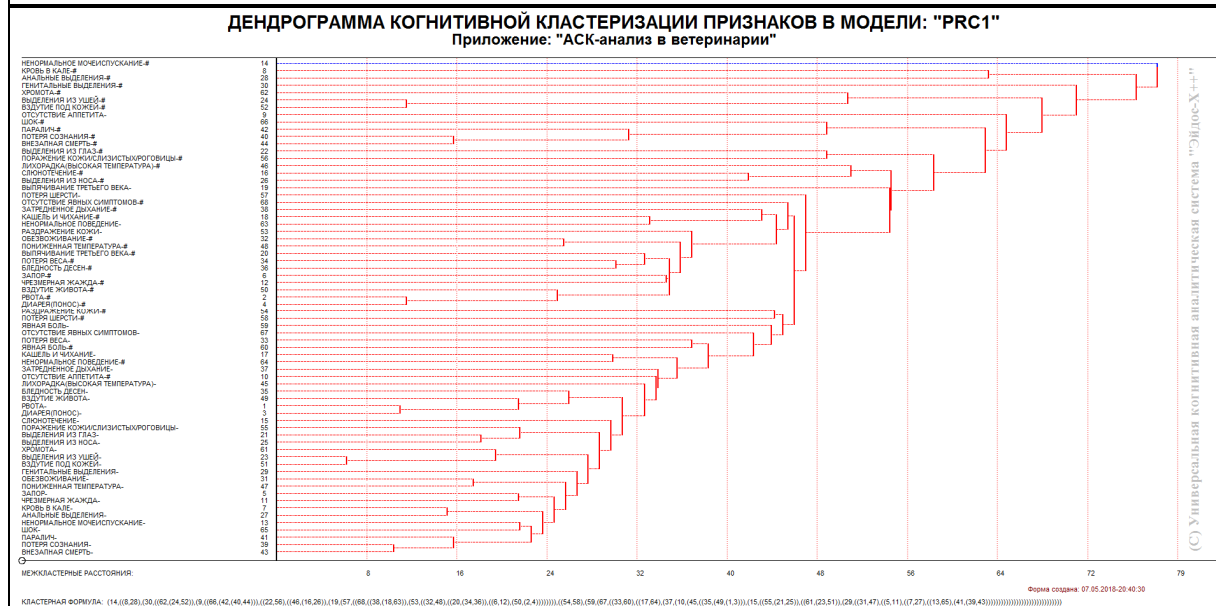
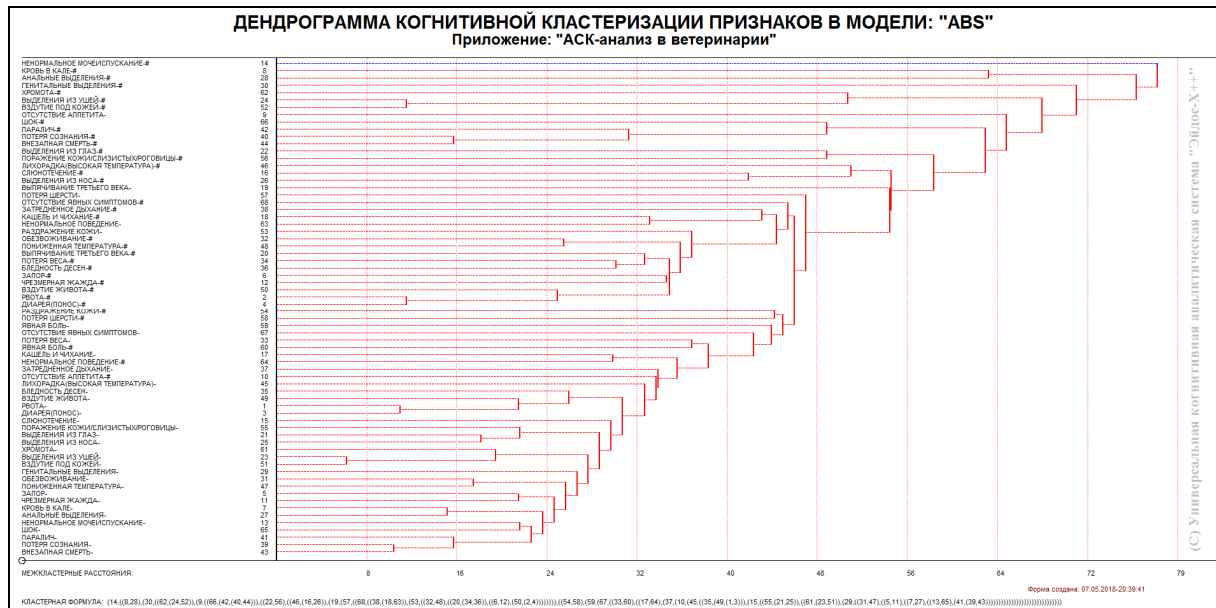
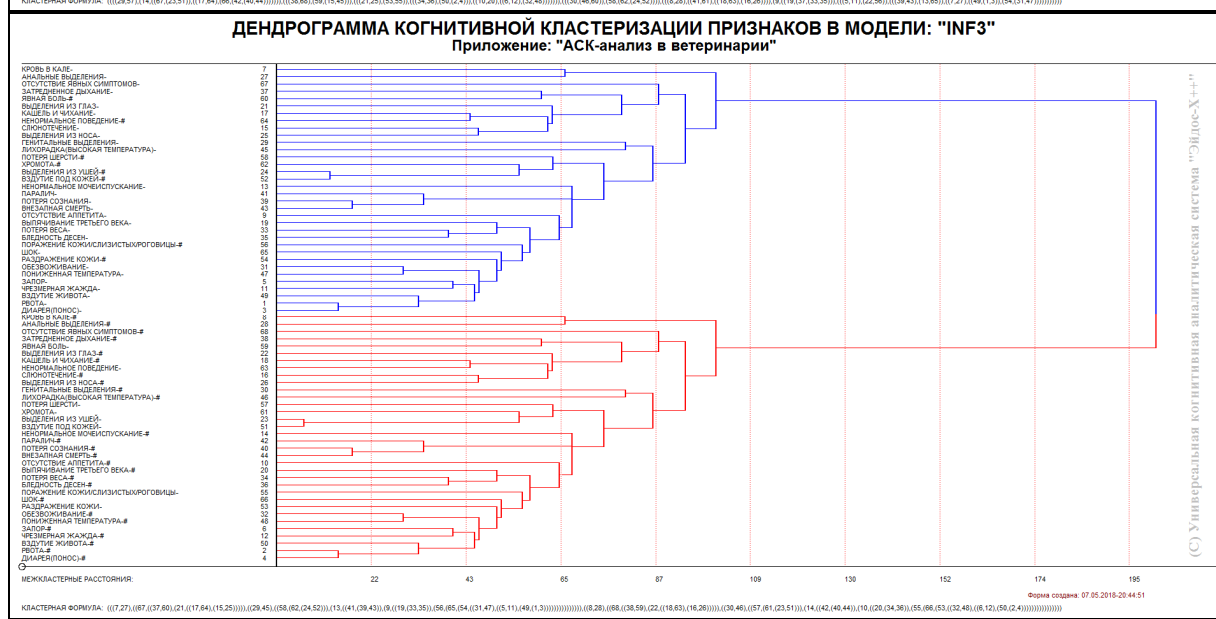
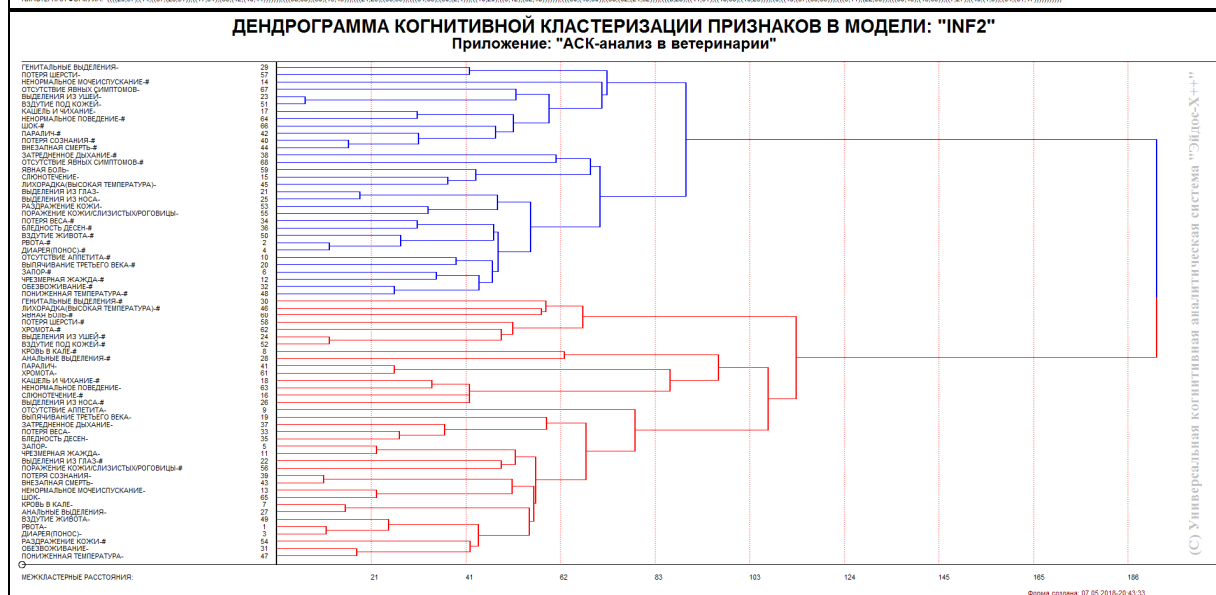
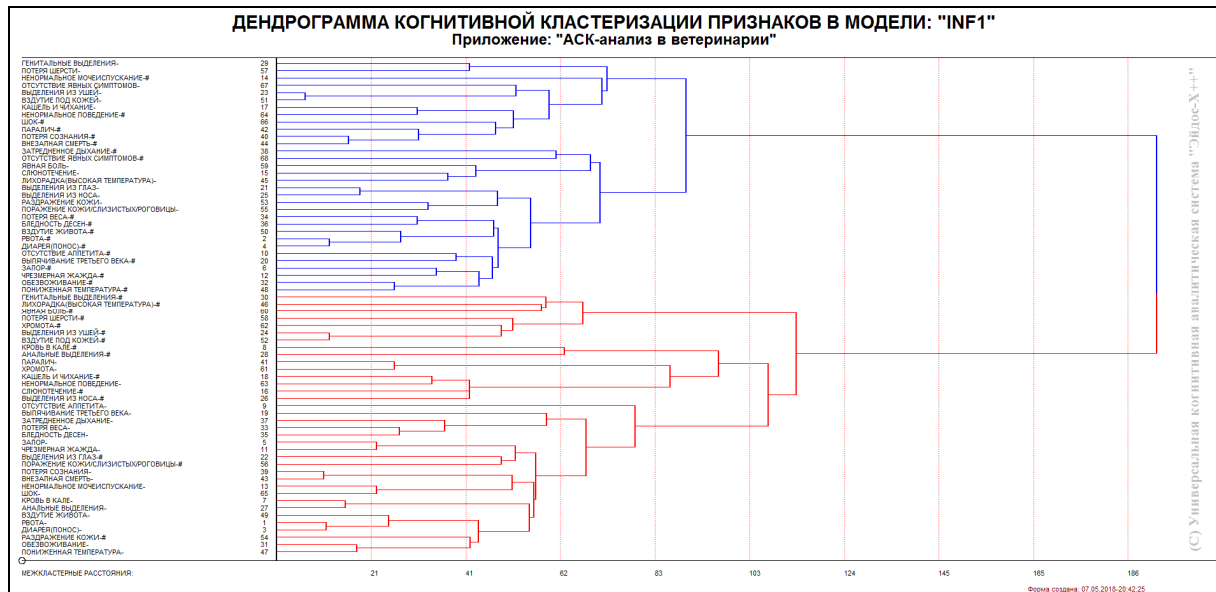
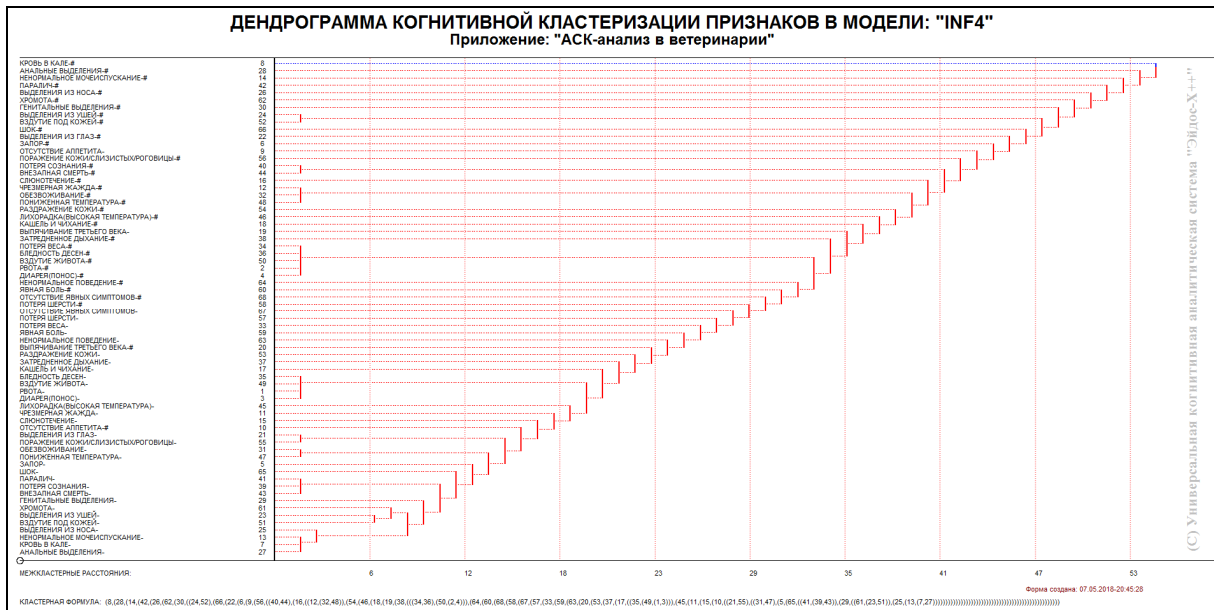


Рисунок 1. Help режима агломеративной когнитивной кластеризации признаков

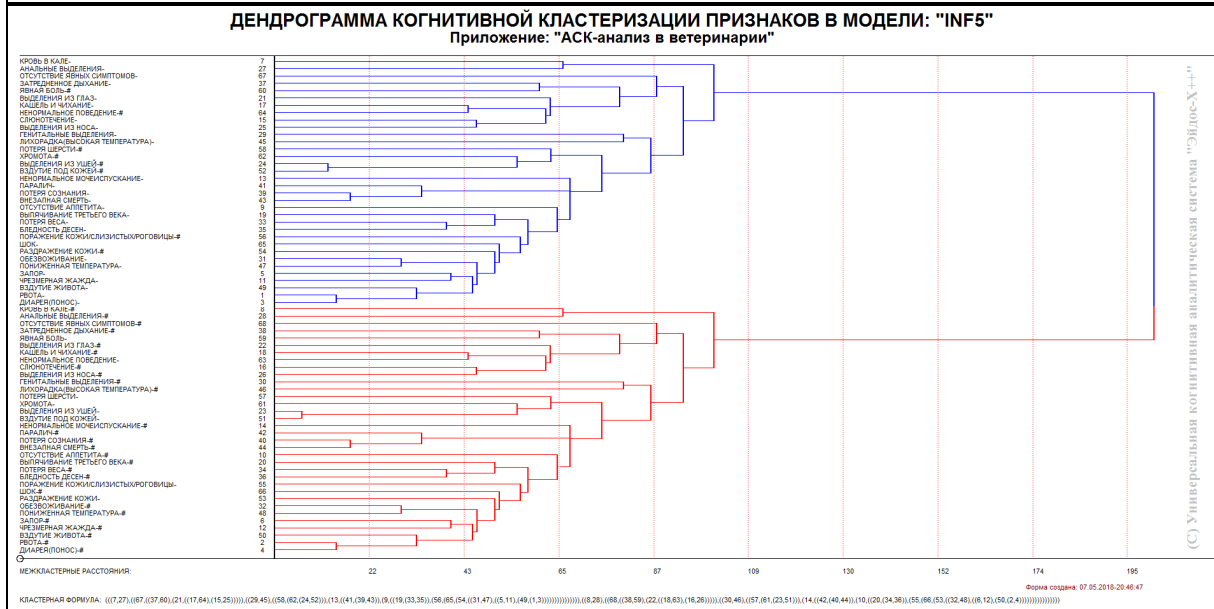
На рисунках 2 приведены агломеративные дендрограммы симптомов и синдромов. При этом под синдромом, как мы уже говорили выше, мы будем понимать несколько симптомов, объединенных в группу по сходству их диагностического смысла.



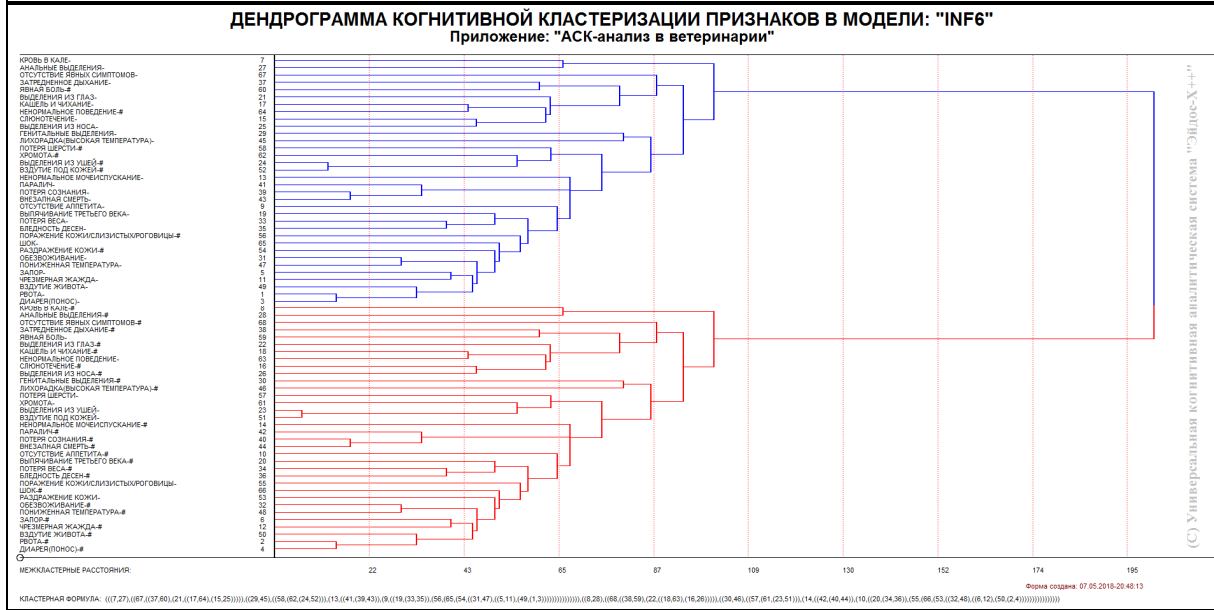




(С) Универсальная когнитивная аналитическая система "Эйдос-X++"



(С) Универсальная когнитивная аналитическая система "Эйдос-X++"



(С) Универсальная когнитивная аналитическая система "Эйдос-X++"

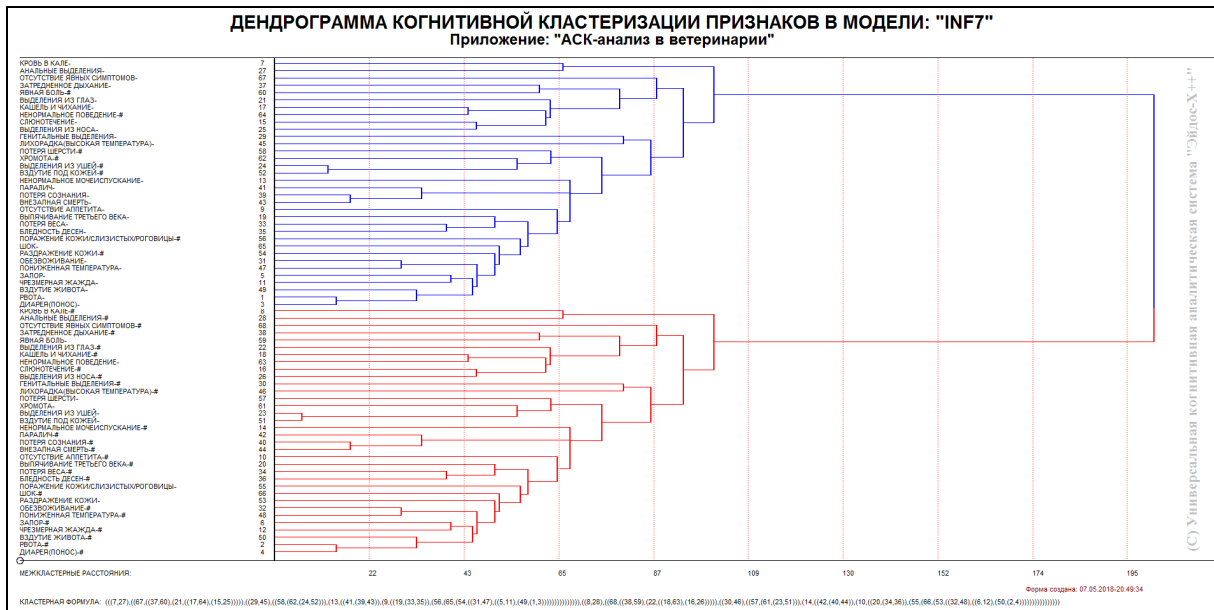
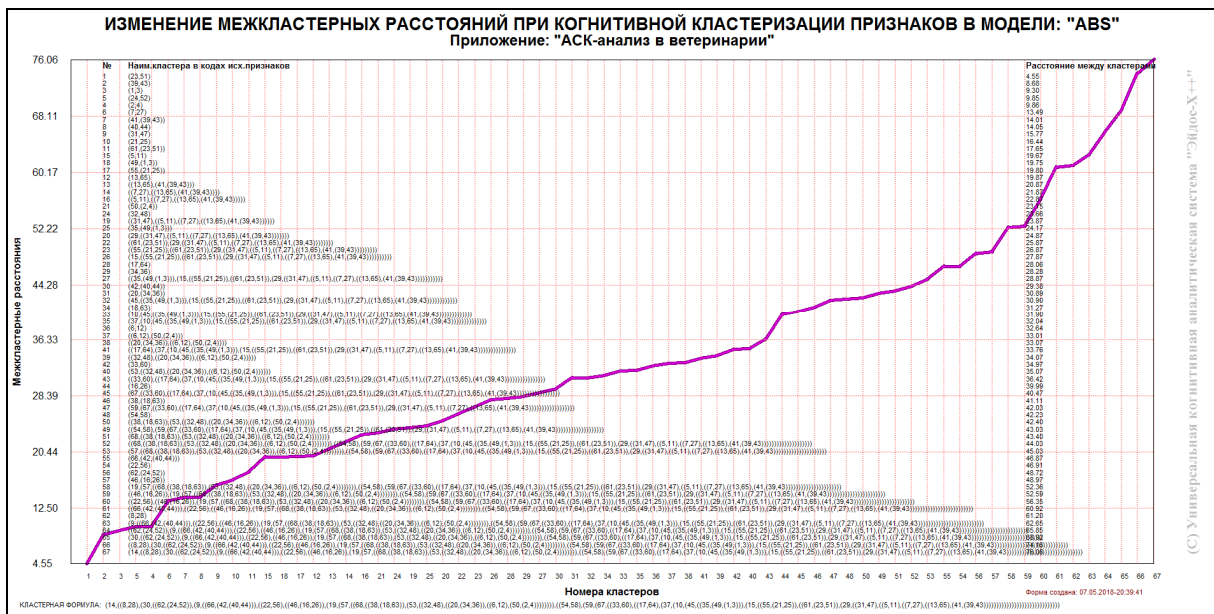
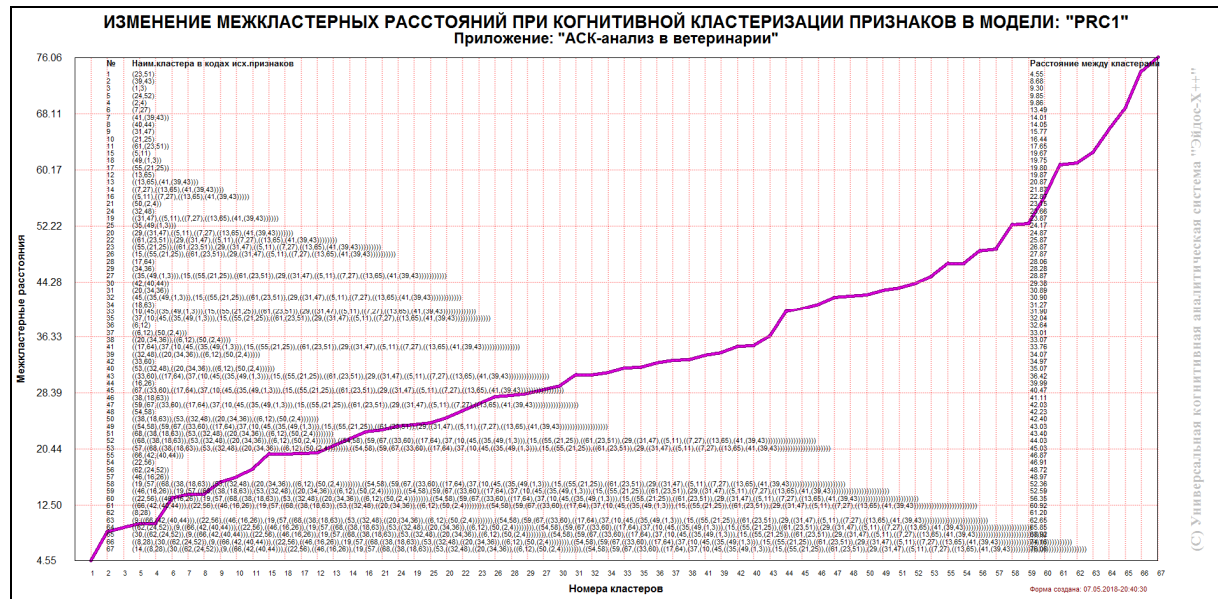


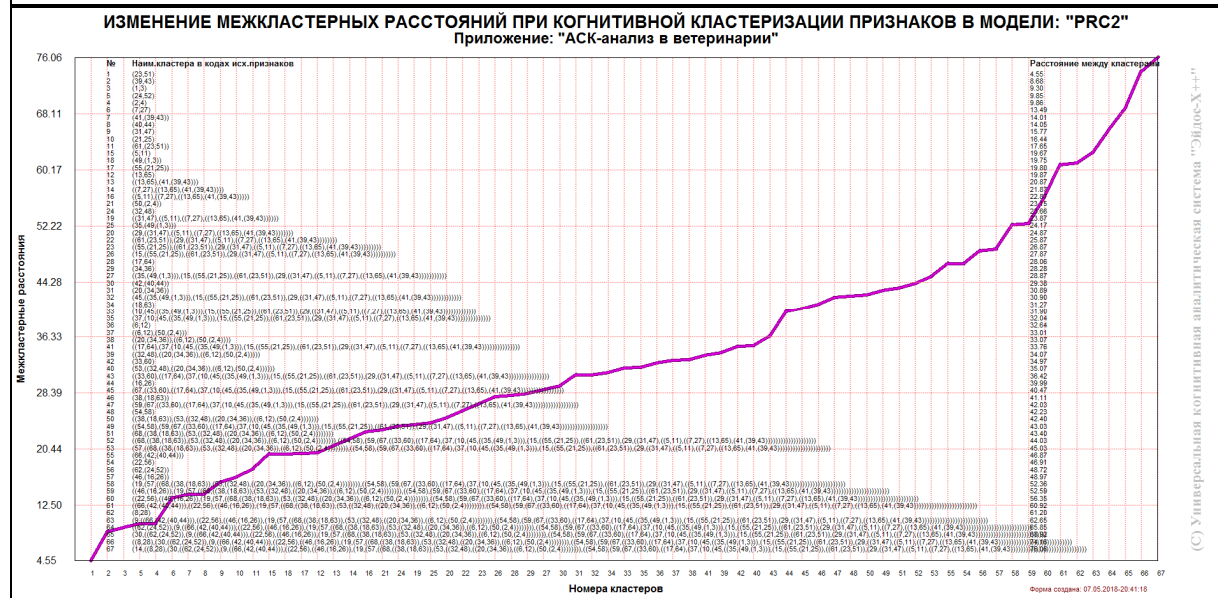
Рисунок 2. Дендрограммы симптомов и синдромов, построенные на основе моделей, созданных и исследованных в работах [5, 6, 7]

На рисунке 3 приведены графики межкластерных расстояний симптомов и синдромов, построенные на основе моделей, созданных и исследованных в работах [5, 6, 7].

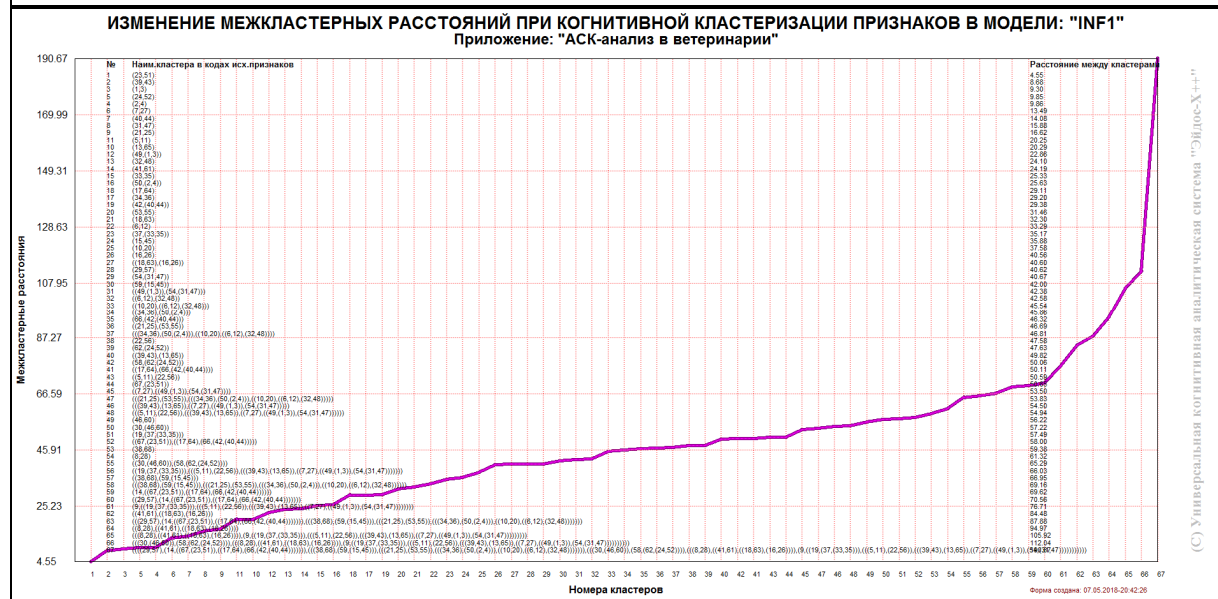




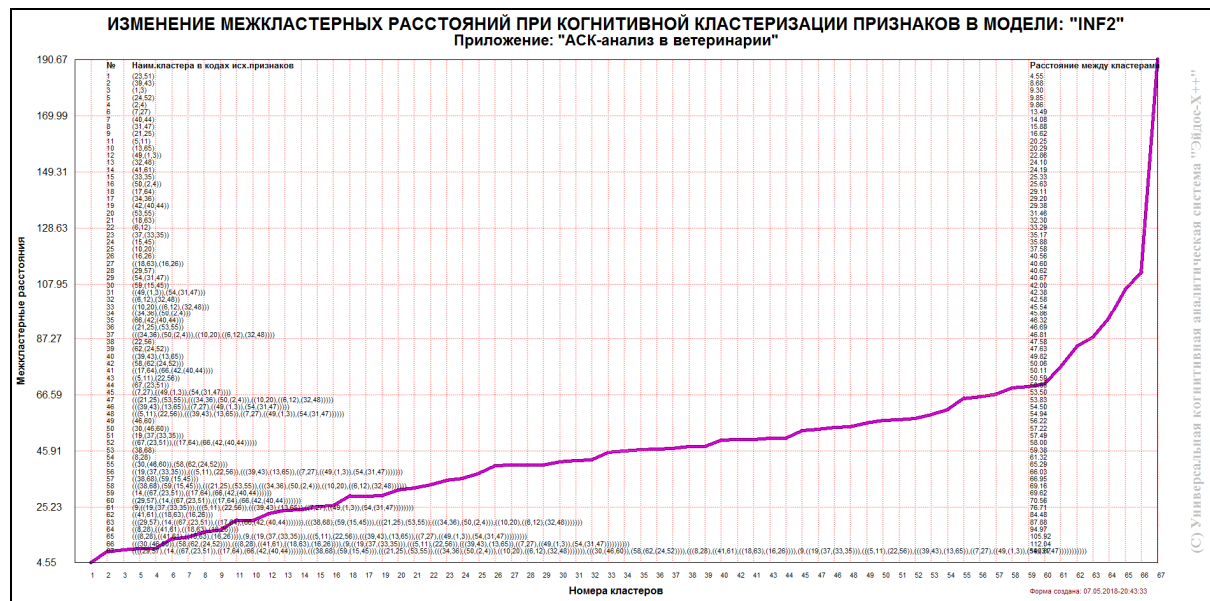
© Универсальная когнитивная аналитическая система "Эйдос-X++"



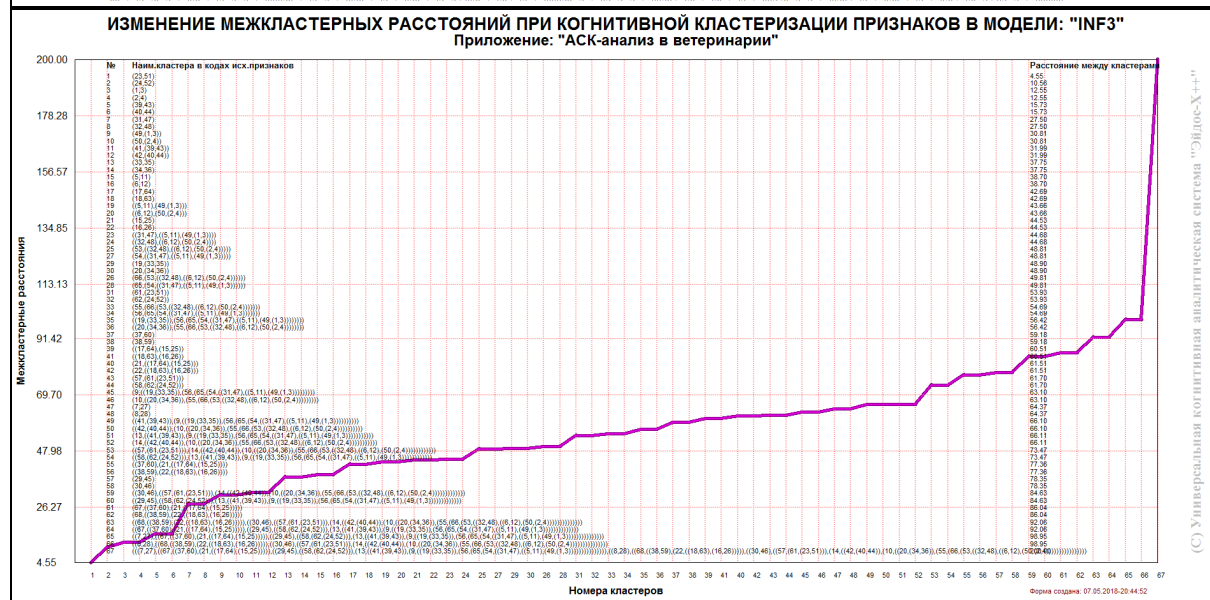
© Универсальная когнитивная аналитическая система "Эйдос-X++"



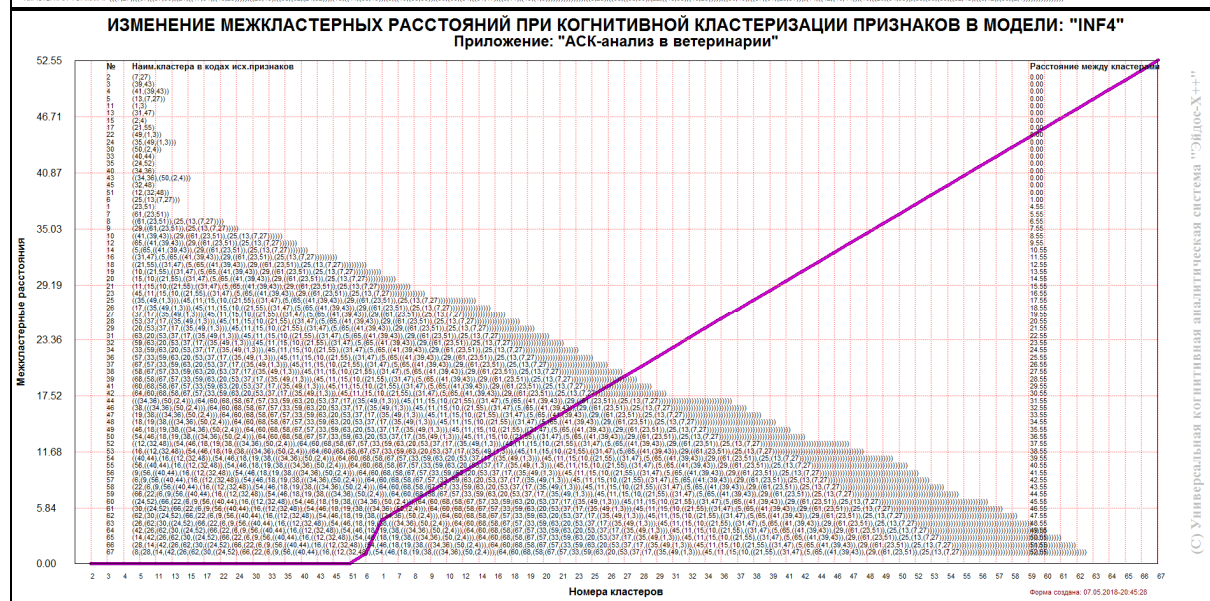
© Универсальная когнитивная аналитическая система "Эйдос-X++"



© Универсальная когнитивная аналитическая система "Эйдос-X++"



© Универсальная когнитивная аналитическая система "Эйдос-X++"



© Универсальная когнитивная аналитическая система "Эйдос-X++"

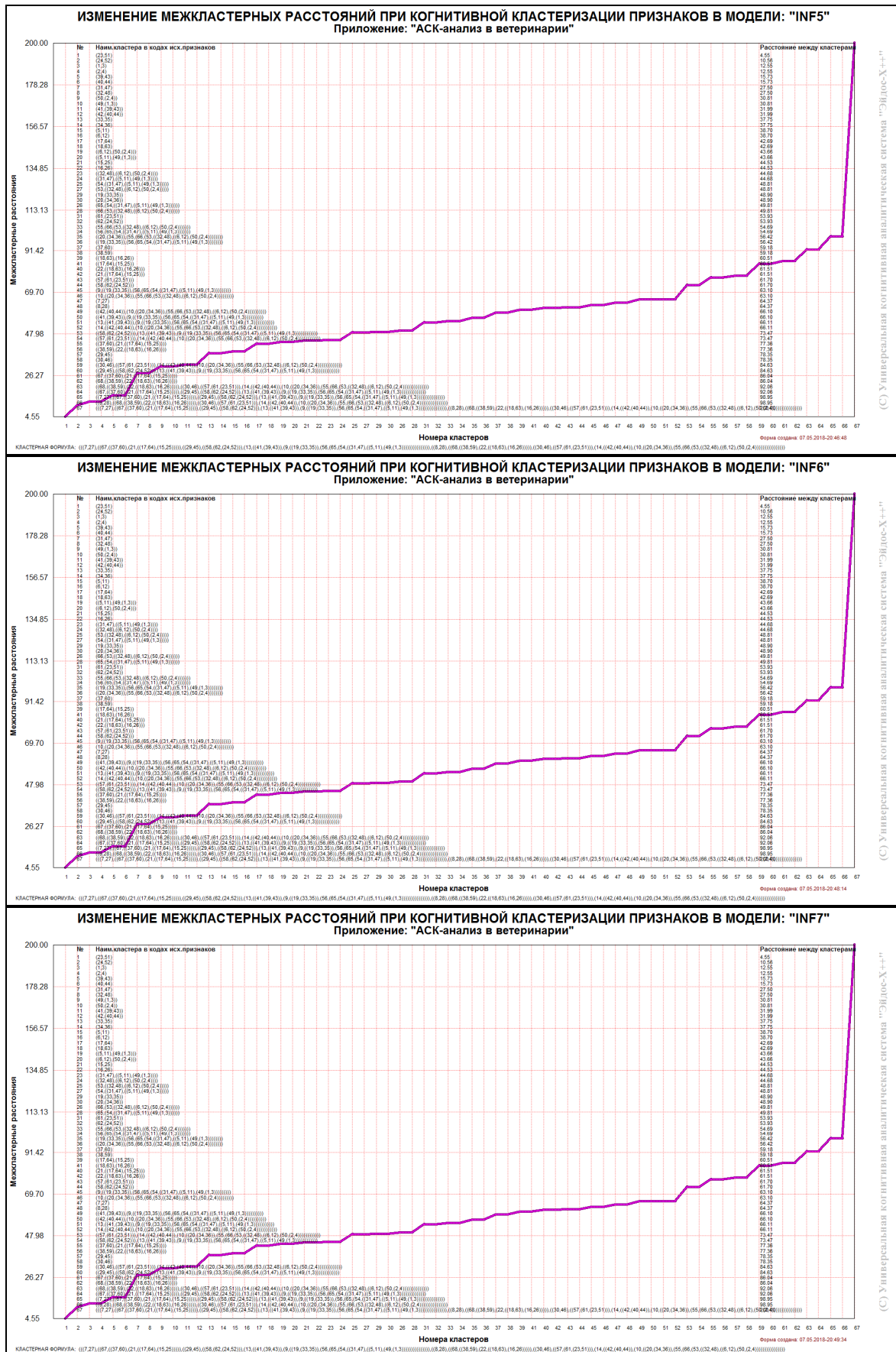


Рисунок 3. Межкластерные расстояния симптомов и синдромов, построенные на основе моделей, созданных и исследованных в работах [5, 6, 7]

Отметим, что в традиционных формах агломеративного кластерного анализа есть *частичный* аналог дендрограммы в модели ABS, а дендрограммы, аналогичные полученным в моделях PRC1, PRC2 и в системно-когнитивных моделях: INF1, INF2, INF3, INF4, INF5, INF6 и INF7 вообще отсутствуют.

Результаты кластерно-конструктивного анализа классов отображаются также в форме когнитивных диаграмм, пример одной из которых приведен на рисунке 4:

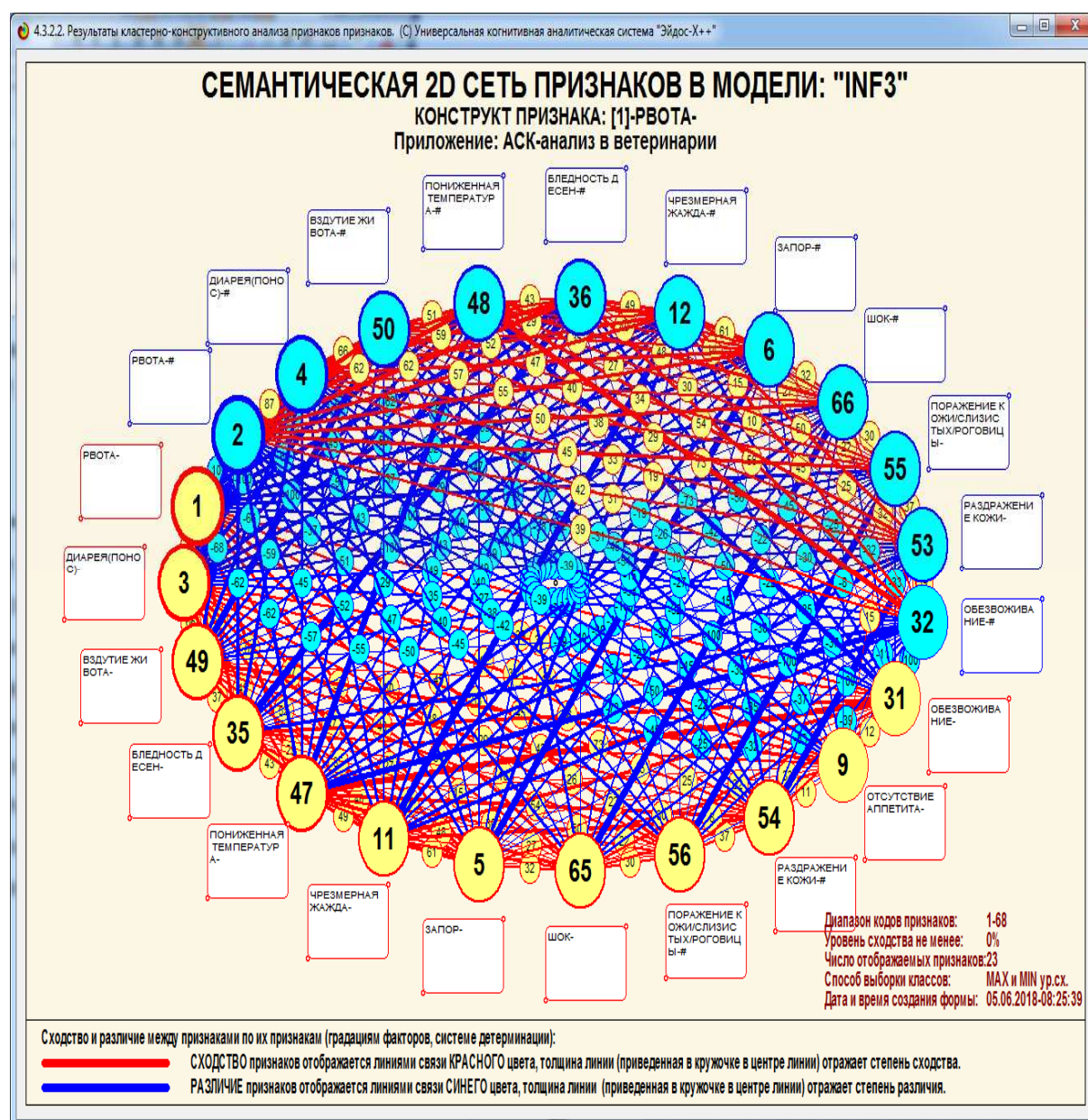


Рисунок 4. Форма управления и когнитивная диаграмма с результатами кластерно-конструктивного анализа симптомов

По результатам статьи можно сделать обоснованный **вывод** о том, что метод «Агломеративной когнитивной кластеризации» является адекватным средством решения поставленной задачи, т.к. позволяет достоверно оценить сходство-различие симптомов и синдромов в ветеринарии по их диагностическому смыслу. Это позволяет ветеринару более обоснованно применять решение о назначении лечения с учетом не только результатов диагностики, но и классификации самих симптомов и синдромов.

В качестве **недостатка** предлагаемого метода можно указать на более значительные затраты времени и вычислительных ресурсов по сравнению с классическими традиционными методами кластеризации.

Как **перспективу** развития метода отметим: повышение размерности дендрограмм и разработку других видов дендрограмм, например, круговых.

Материалы данной статьи могут быть использованы в учебном процессе при преподавании дисциплин: «Интеллектуальные информационные системы», «представление знаний в информационных системах», «Современные технологии в образовании (магистратура)», «Управление знаниями (магистратура)», «Введение в искусственный интеллект», «Интеллектуальные и нейросетевые технологии в образовании (магистратура)», «Основы искусственного интеллекта» и других.

Планируется продолжить исследования и разработки, представленные в данной статье, по следующим направлениям:

1. Агломеративная когнитивная кластеризация нозологических образов, симптомов и синдромов в ветеринарии на конкретных примерах, относящихся к другим животным и группам заболеваний.

2. Разработка в среде интеллектуальной системы Эйдос-Х++ ветеринарных диагностических тестов и проведение научных исследований на основе эмпирических данных (данных наблюдений и экспериментов).

3. Интеграция в среде интеллектуальной системы Эйдос-X++ без программирования нескольких диагностических тестов в батарею тестов или в супертест с одной общей моделью.

Этим и другим применениям способствует и то, что система «Эйдос» является мультязычной интеллектуальной on-line средой для обучения и научных исследований [3, 4]⁵ и находится в полном открытом бесплатном доступе (причем с подробно комментированными открытыми исходными текстами: http://lc.kubagro.ru/_AIDOS-X.txt) на сайте автора по адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm.

В заключение автор выражает благодарность проректору по научной работе Кубанского ГАУ им. И.Т. Трубилина доктору биологических наук профессору Андрею Георгиевичу Коцаеву за помощь в публикации статьи.

Список литературы

1. 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.

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

3. Луценко Е.В. Открытая масштабируемая интерактивная интеллектуальная on-line 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.

4. Луценко Е.В., Открытая масштабируемая интерактивная интеллектуальная on-line 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.

5. Луценко Е.В. Автоматизированный системно-когнитивный анализ в ветеринарии (на примере разработки диагностических тестов) / Е.В. Луценко // Политематиче-

⁵ http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

ский сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №03(137). С. 143 – 196. – IDA [article ID]: 1371803031. – Режим доступа: <http://ej.kubagro.ru/2018/03/pdf/31.pdf>, 3,375 у.п.л.

6. Луценко Е.В. Агломеративная когнитивная кластеризация нозологических образцов в ветеринарии / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №04(138). С. 122 – 139. – IDA [article ID]: 1381804033. – Режим доступа: <http://ej.kubagro.ru/2018/04/pdf/33.pdf>, 1,125 у.п.л.

7. Луценко Е.В. Реализация тестов и супертестов для ветеринарной и медицинской диагностики в среде системы искусственного интеллекта «Эйдос-Х++» без программирования / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №05(089). С. 167 – 207. – IDA [article ID]: 0891305014. – Режим доступа: <http://ej.kubagro.ru/2013/05/pdf/14.pdf>, 2,562 у.п.л.

References

1. Lucenko E.V. Metod kognitivnoj klasterizacii ili klasterizaciya na osnove znanij (klasterizaciya v sistemno-kognitivnom analize i intellektual`noj sisteme «E`jdos») / E.V. Lucenko, V.E. Korzhakov // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2011. – №07(071). S. 528 – 576. – Shifr Informregistra: 0421100012\0253, IDA [article ID]: 0711107040. – Rezhim dostupa: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>, 3,062 у.п.л.

2. Lucenko E.V., Podсистема aglomerativnoj kognitivnoj klasterizacii klassov sistemy` «E`jdos» ("E`jdos-klaster"). Pat. № 2012610135 RF. Zayav. № 2011617962 RF 26.10.2011. Opubl. Ot 10.01.2012. – Rezhim dostupa: <http://lc.kubagro.ru/aidos/2012610135.jpg>, 3,125 у.п.л.

3. Lucenko E.V. Otkry`taya masshtabiruemaya interaktivnaya intellektual`naya on-line sreda dlya obucheniya i nauchny`x issledovanij na baze ASK-analiza i sistemy` «E`jdos» / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2017. – №06(130). S. 1 – 55. – IDA [article ID]: 1301706001. – Rezhim dostupa: <http://ej.kubagro.ru/2017/06/pdf/01.pdf>, 3,438 у.п.л.

4. Lucenko E.V., Otkry`taya masshtabiruemaya interaktivnaya intellektual`naya on-line sreda «E`jdos» («E`jdos-online»). Svid. RosPatenta RF na programmu dlya E`VM, Zayavka № 2017618053 ot 07.08.2017, Gos.reg.№ 2017661153, zaregistr. 04.10.2017. – Rezhim dostupa: <http://lc.kubagro.ru/aidos/2017661153.jpg>, 2 у.п.л.

5. Lucenko E.V. Avtomatizirovanny`j sistemno-kognitivny`j analiz v veterinarii (na primere razrabotki diagnosticheskix testov) / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2018. – №03(137). S. 143 – 196. – IDA [article ID]: 1371803031. – Rezhim dostupa: <http://ej.kubagro.ru/2018/03/pdf/31.pdf>, 3,375 у.п.л.

6. Lucenko E.V. Aglomerativnaya kognitivnaya klasterizaciya nozologicheskix obrazov v veterinarii / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU)

[E`lektronny`j resurs]. – Krasnodar: KubGAU, 2018. – №04(138). S. 122 – 139. – IDA [article ID]: 1381804033. – Rezhim dostupa: <http://ej.kubagro.ru/2018/04/pdf/33.pdf>, 1,125 u.p.l.

7. Lucenko E.V. Realizaciya testov i supertestov dlya veterinarnoj i medicinskoj diagnostiki v srede sistemy` iskusstvennogo intellekta «E`jdos-X++» bez programmirovaniya / E.V. Lucenko, V.E. Korzhakov // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2013. – №05(089). S. 167 – 207. – IDA [article ID]: 0891305014. – Rezhim dostupa: <http://ej.kubagro.ru/2013/05/pdf/14.pdf>, 2,562 u.p.l.