

УДК 004.8 (075.8)
ББК 32.965
JEL Code C45, C87, C88, D83, D87

Автоматизированный системно-когнитивный анализ влияния выбора мест приобретения крупного рогатого скота на объем и структуру его заболеваемости

Lutsenko Eugeny Veniaminovich¹, Semenenko Kseniya Andreevna¹,

¹Kuban state agrarian University named after I. T. Trubilin, Russia
prof.lutsenko@gmail.com, <http://lc.kubagro.ru>, ksenia.semenenko@gmail.com

Аннотация: В практике выращивания крупного рогатого скота мясного и молочного направления в Краснодарском крае сложилась устойчивая многолетняя традиция ежегодно закупать большое количество голов из-за пределов края и из-за рубежа. Понятно, что это делается для улучшения натуральных и финансово-экономических показателей работы мясомолочной отрасли. Однако на практике достижению этих целей препятствует заболеваемость крупного рогатого скота. С этой заболеваемостью, когда она является уже состоявшимся фактом, борются ветеринарные службы. Но существуют факторы, снижающие риск самой заболеваемости. Например, существуют ярко-выраженные причинно-следственные зависимости объемов и структуры заболеваемости крупного рогатого скота от мест его закупки. Однако в мире и России практически нет исследований, специально посвященных изучению этих зависимостей. Данная работа призвана восполнить этот существенный пробел и этим обусловлена ее актуальность. Целью работы является количественное изучение зависимости объемов и структуры заболеваемости крупного рогатого скота от мест его закупки. Для выявления и изучения этих зависимостей на основе непосредственно эмпирических данных был применен новый инновационный метод искусственного интеллекта Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос». В результате проведенного исследования были созданы статистические и системно-когнитивные модели, имеющие очень высокую достоверность по F-мере Ван Ризбергера, обобщенной автором. На основе этих моделей получены графические и табличные выходные формы, отражающие систему детерминации, обуславливающую высокий, средний и низкий уровни заболеваемости крупного рогатого скота различными заболеваниями. Основной вывод, который можно обоснованно сделать по результатам работы, состоит в том, что для снижения заболеваемости крупного рогатого скота нужно минимизировать его закупки из-за пределов региона и особенно из таких стран как Австрия и Германия. Авторы выражают надежду, что данная статья будет полезна специалистам, в область компетенции которых входит принятие решений и закупках крупного рогатого скота вне региона и из других стран.

Ключевые слова: атрибуция текстов Автоматизированный системно-когнитивный анализ, интеллектуальная система «Эйдос».

1. Вступление (Introduction)

В данной работе предлагается новая математическая модель, основанная на теории информации, обеспечивающая выявление силы и направления мест его закупки крупного рогатого скота на объемов и структуру его заболеваемости. Кроме того разработаны соответствующая методика численных расчетов и реализующий их программный инструментарий.

Цель работы является количественное изучение зависимости объемов и структуры заболеваемости крупного рогатого скота от мест его закупки. Для достижения этой цели были поставлены и решены следующие задачи:

Задача-1: формирование обобщенных образов классов по объемам и структуре заболеваемости крупного рогатого скота на основе многолетних данных строгой отчетности региональных ветеринарных служб;

Задача-2: количественное сравнение образов конкретных лет с обобщенными образами групп (классов);

Задача-3: сравнение обобщенных образов классов друг с другом и разработка кластеров и конструктов;

Задача-4: исследование моделируемой предметной области путем исследования ее системно-когнитивной модели.

Все эти задачи решаются с помощью авторского программного и научно-методического инструментария, представляющего собой открытое программное обеспечение.

2. Обзор литературы (Literature review)

С появлением и развитием интеллектуальных технологий область их успешного применения постоянно и быстро расширяется. Научные исследования и разработки в этой важной области традиционно посвящены в основном разработке концептуальных подходов и математических моделей, тогда как для конкретных исследователей не столь важно иметь детальное описание математической модели, сколько реализующий ее программный инструментарий, эффективный для решения конкретных задач.

Однако, судя по литературным данным, интеллектуальные технологии практически не применялись для анализа последствий закупок крупного рогатого скота у различных нерегиональных и зарубежных поставщиков. По крайней мере, в англоязычных поисковых системах как это ни странно не удалось найти работ, посвященных этой тематике.

И это не случайно. Дело в том, что путь от математической модели до реализующей ее программной системы долог и тернист. Этот путь включает разработку методики численных расчетов (т.е. структур данных и алгоритмов их обработки), а также программной реализации. Все это предполагает значительные затраты различных видов ресурсов: прежде всего интеллектуальных ресурсов, но также и длительного времени, и хорошего финансирования. Судя по всему в настоящее время интеллектуальные технологии еще слишком дороги и сложны в применении для ветеринаров. Авторы в меру своих сил пытались решить эту проблему или по крайней мере подготовить почву для ее решения [1-11].

В данной статье приводится новая, ранее не описанная в мировой литературе, математическая модель, основанная на теории информации, а также реализующий ее программный инструментарий и описывается их применение для решения поставленных задач. Принципиально важно отметить, что этот программный инструментарий не предполагает у пользователя предварительных компетенций в области искусственного интеллекта и является открытым программным обеспечением, т.е. имеет практически «нулевой порог входа».

В этом и заключается актуальность статьи и ее вклад в российскую и мировую науку и практику выращивания крупного рогатого скота.

2. Материалы и методы (Materials and methods)

2.1. Обоснование выбора метода и инструментария решения проблемы

As a method of the research we used Automated system-cognitive analysis (ASC-analysis), which is a new innovative method of artificial intelligence: it also has its own software tool – an intelligent system called "Eidos" (open source software) [1-11].

The Eidos-X++ system differs from other artificial intelligence systems in the following parameters:

- was developed in a universal setting, independent of the subject area. Therefore, it is universal and can be applied in many subject areas (<http://lc.kubagro.ru/aidos/index.htm>);
- is in full open free access (<http://lc.kubagro.ru/aidos/Aidos-X.htm>), and with the relevant source texts (<http://lc.kubagro.ru/AIDOS-X.txt>);
- is one of the first domestic systems of artificial intelligence of the personal level, i.e. it does not take special training in the field of technologies of artificial intelligence from the user (there is an act of introduction of system "Eidos" of 1987) (<http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>);
- provides stable identification in a comparable form of strength and direction of cause-effect relationships in incomplete noisy interdependent (nonlinear) data of very large dimension of numerical and non-numerical nature, measured in different types of scales (nominal, ordinal and numerical) and in different units of measurement (i.e. does not impose strict requirements to the data that can not be performed, and processes the data that is);
- contains a large number of local (supplied with the installation) and cloud educational and scientific applications (currently 31 and 201 (http://aidos.byethost5.com/Source_data_applications/WebAppls.htm), respectively) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);
- provides multilingual interface support in 51 languages. Language databases are included in the installation and can be replenished automatically;
- supports on-line environment of knowledge accumulation and is widely used all over the world (<http://aidos.byethost5.com/map5.php>);
- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решения этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний;

- плохой перевод: the most time-consuming computationally, the operations of the synthesis models and implements recognition by using graphic processing unit (GPU) that some tasks can only support up to the solution of these tasks is several thousand times that really provides intelligent processing of big data, big information and big knowledge;

- provides transformation of the initial empirical data into information, and its knowledge and solution using this knowledge of classification problems, decision support and research of the subject area by studying its system-cognitive model, generating a very large number of tabular and graphical output forms (development of cognitive graphics), many of which have no analogues in other systems (examples of forms can be found in: http://lc.kubagro.ru/aidos/aidos18_LLS/aidos18_LLS.pdf);

- well imitates the human style of thinking: gives the results of the analysis, understandable to experts on the basis of their experience, intuition and professional competence.

Этапы АСК-анализа:

1. Когнитивно-целевая структуризация предметной области:

- разработка классификационных шкал;

- разработка описательных шкал.

2. Формализация предметной области:

- разработка градаций классификационных шкал;

- разработка градаций описательных шкал.

- кодирование исходных данных с помощью классификационных и описательных шкал и градаций и формирование обучающей выборки.

3. Синтез, повышение качества и верификация статистических и системно-когнитивных моделей.

4. Решение в наиболее достоверной модели задач:

- идентификации (распознавания, диагностики, классификации);

- поддержки принятия решений;

- исследования моделируемой предметной области путем исследования ее модели.

Рассмотрим суть математической модели АСК-анализа и ее применение для решения поставленных задач.

2.2. Суть метода и математической модели АСК-анализа

Суть метода АСК-анализа состоит в последовательном повышении степени формализации модели и преобразовании данных в информацию, а ее в знания и решении на основе этих знаний задач идентификации (распознавания, классификации и прогнозирования), поддержки принятия решений и исследования моделируемой предметной области.

В АСК-анализе все значения факторов рассматриваются *с одной единственной точки зрения: сколько информации содержится в их значениях о переходе объекта моделирования и управления, на который они действуют, в определенное будущее состояние, описываемое классом (градация классификационной шкалы), и при этом сила и направление влияния всех значений факторов на объект измеряется в одних общих для всех факторов единицах измерения: единицах количества информации.*

2.3. Синтез системно-когнитивных моделей и частные критерии знаний, многопараметрическая типизация

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике и обеспечивает сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых данных, представленных в различных типах шкал (номинальных, порядковых и числовых) и различных единицах измерения.

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных рассчитывается матрица абсолютных частот. Используя «программистский стиль» расчет этой матрицы записывается с помощью формул (1). Если i -е слово обнаружено в тексте, принадлежащем j -му классу, то:

$$N_{ij} = N_{ij} + 1; N_{i\Sigma} = N_{i\Sigma} + 1; N_{\Sigma j} = N_{\Sigma j} + 1; N_{\Sigma\Sigma} = N_{\Sigma\Sigma} + 1 \quad (1)$$

На основе выражений (1) по формулам (2) рассчитываются матрицы условных и безусловных процентных распределений:

$$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}, P_{i\Sigma} = \frac{N_{i\Sigma}}{N_{\Sigma\Sigma}} \quad (2)$$

Отметим, что в АСК-анализе и его программном инструментарии интеллектуальной системе «Эйдос» используется два способа расчета матриц условных и безусловных процентных распределений:

1-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество признаков по классу;

2-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество объектов обучающей выборки по классу.

Затем на основе выражений (2) с использованием частных критериев, приведенных формулах (3), рассчитываются матрицы системно-когнитивных моделей:

$$I_{ij} = \Psi \times \log_2 \frac{P_{ij}}{P_i}, \quad I_{ij} = \Psi \times \log_2 \frac{N_{ij}N}{N_i N_j}, \quad I_{ij} = N_{ij} - \frac{N_i N_j}{N}, \quad I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$$

$$I_{ij} = \frac{N_{ij}N}{N_i N_j} - 1, \quad I_{ij} = P_{ij} - P_i, \quad I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$$

$$\sigma_{ix} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}, \quad H = \sqrt{\frac{1}{(W \cdot M - 1) \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}}$$

Обозначения к формулам (1)-(3):

i – значение фактора; j – класс; N_{ij} – количество встреч i -го значения фактора в наблюдениях объектов обучающей выборки j -го класса; N_i – количество встреч i -го значения фактора по всей обучающей выборке; N_j – количество значений факторов в наблюдениях объектов j -го класса; N – количество значений факторов у всех объектов обучающей выборки; I_{ij} – частный критерий знаний: количество знаний в факте наблюдения i -го значения фактора о том, что объект, на который действует это значение фактора относится к j -му классу; Ψ – нормировочный коэффициент (Е.В.Луценко, 2002), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли; P_i – безусловная относительная частота встречи i -го значения фактора в обучающей выборке; P_{ij} – условная относительная частота встречи i -го значения фактора у объектов j -го класса; W – количество классов; M – количество различных значений факторов во всех объектах обучающей выборки; σ_i – ценность i -го значения фактора для классификации наблюдений; H – качество модели..

На основе системно-когнитивных моделей, отличаются частыми критериями (3), решаются задачи идентификации (классификации, распознавания, диагностики, прогнозирования), поддержки принятия решений (обратная задача прогнозирования), а также задача исследования моделируемой предметной области путем исследования ее системно-когнитивной модели.

Для решения этих задач в АСК-анализе и системе «Эйдос» в настоящее время используется два аддитивных интегральных критерия.

2.4. Интегральные критерии и решение задач системной идентификации и принятия решений

Задача системной идентификации – это задача определения степени сходства (и различия) конкретного текста с обобщенными образами классов, соответствующих будущим состояниям объекта моделирования. В моделях, приведенных в выражениях (4), отражено, какое количество информации содержится в каждом значении фактора о принадлежности объекта моделирования к каждому из классов. Но в на объект моделирования действует много значений факторов.

Поэтому естественно считать, что объект моделирования под действием этих факторов перейдет в будущие состояния, соответствующие тем классам, о переходе к которым в этих значениях факторов содержится максимальное **суммарное** количество информации.

Функция от частных критериев, имеющая определенное числовое значение, свое для каждого класса и отражающее степень принадлежности объекта моделирования к данному классу, называется **интегральным критерием**.

В результате получается, что некоторый определенный конкретный объект в различной степени принадлежит и не принадлежит к разным классам. Интегральные критерии применяются при решении как задачи идентификации или прогнозирования, так и задачи принятия решений. В настоящее время в системе «Эйдос» используется два аддитивных интегральных критерия: сумма знаний и резонанс знаний.

1-й интегральный критерий «Сумма знаний» представляет собой суммарное количество знаний, содержащееся в признаках объекта о его сходстве/различии с каждым из классов. Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний(4):

$$I_j = (\vec{I}_{ij}, \vec{L}_i). \quad (4)$$

В выражении круглыми скобками обозначено скалярное произведение. В координатной форме это выражение имеет вид (5):

$$I_j = \sum_{i=1}^M I_{ij} L_i, \quad (5),$$

где:

$\bar{I}_{ij} = \{I_{ij}\}$ – вектор j -го класса; $\bar{L}_i = \{L_i\}$ – вектор распознаваемого текста, т.е.:

$$\bar{L}_i = \begin{cases} 1, & \text{если } i - \text{й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i - \text{й фактор действует с истинностью } n; \\ 0, & \text{если } i - \text{й фактор не действует.} \end{cases}$$

2-й интегральный критерий «Семантический резонанс знаний» представляет собой нормированное суммарное количество знаний, содержащееся в признаках объекта о его сходстве/различии с каждым из классов. Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний и имеет вид (6):

$$I_j = \frac{1}{\sigma_j \sigma_i M} \sum_{i=1}^M (I_{ij} - \bar{I}_j) (L_i - \bar{L}), \quad (6)$$

где:

σ_j – среднее квадратичное отклонение частных критериев знаний вектора класса;

σ_i – среднее квадратичное отклонение по вектору распознаваемого текста.

3. Результаты (Data, Analysis, and Results)

3.1. Когнитивно-целевая структуризация предметной области

Когнитивно-целевая структуризация предметной области включает разработку классификационных и описательных шкал (таблицы 1 и 2).

Таблица 1. Классификационные шкалы

Код	Наименование
1	TOTAL NUMBER OF CASES: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
2	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
3	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (%)
4	TOTAL NUMBER OF CASES: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
5	TOTAL NUMBER OF CASES: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%)
6	INCLUDING: - DISEASES OF THE DIGESTIVE SYSTEM; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
7	INCLUDING: - DISEASES OF THE DIGESTIVE SYSTEM; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (%)
8	INCLUDING: - DISEASES OF THE DIGESTIVE SYSTEM; - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
9	INCLUDING: - DISEASES OF THE DIGESTIVE SYSTEM; - DIED CATTLE, INCL. YOUNG CATTLE (%)
10	INCLUDING: - DISEASES OF THE DIGESTIVE SYSTEM; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
11	INCLUDING: - DISEASES OF THE DIGESTIVE SYSTEM; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%)
12	- DISEASES OF THE RESPIRATORY SYSTEM; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
13	- DISEASES OF THE RESPIRATORY SYSTEM; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (%)
14	- DISEASES OF THE RESPIRATORY SYSTEM; - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
15	- DISEASES OF THE RESPIRATORY SYSTEM; - DIED CATTLE, INCL. YOUNG CATTLE (%)
16	- DISEASES OF THE RESPIRATORY SYSTEM; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
17	- DISEASES OF THE RESPIRATORY SYSTEM; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%)
18	- METABOLIC DISEASES; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
19	- METABOLIC DISEASES; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (%)
20	- METABOLIC DISEASES; - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
21	- METABOLIC DISEASES; - DIED CATTLE, INCL. YOUNG CATTLE (%)
22	- METABOLIC DISEASES; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
23	- METABOLIC DISEASES; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%)
24	- DISEASES OF THE REPRODUCTIVE ORGANS: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
25	- DISEASES OF THE REPRODUCTIVE ORGANS: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (%)
26	- DISEASES OF THE REPRODUCTIVE ORGANS: - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
27	- DISEASES OF THE REPRODUCTIVE ORGANS: - DIED CATTLE, INCL. YOUNG CATTLE (%)
28	- DISEASES OF THE REPRODUCTIVE ORGANS: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
29	- DISEASES OF THE REPRODUCTIVE ORGANS: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%)
30	- INCLUDING MASTITIS; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
31	- INCLUDING MASTITIS; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (%)
32	- INCLUDING MASTITIS; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
33	- INCLUDING MASTITIS; - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%)
34	- POISONING; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
35	- POISONING; - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (%)
36	- POISONING; - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
37	- POISONING; - DIED CATTLE, INCL. YOUNG CATTLE (%)
38	- INJURIES: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)
39	- INJURIES: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (%)
40	- INJURIES: - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
41	- INJURIES: - DIED CATTLE, INCL. YOUNG CATTLE (%)
42	- INJURIES: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)
43	- INJURIES: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%)

Когнитивно-целевая структуризация предметной области – это единственный неавтоматизированный этап АСК-анализа. Классификационные шкалы описывают будущие состояния объекта моделирования, в нашем случае – это объемы и структура заболеваемости. Описательные шкалы описывают факторы, влияющие на объект моделирования, в данной работе – это места и объемы приобретения крупного рогатого скота.

Таблица 2. Описательные шкалы

Код	Наименование
1	HUNGARY
2	AUSTRIA
3	USA
4	NETHERLANDS
5	DENMARK
6	GERMANY
7	TOTAL FROM ALL COUNTRIES
8	CATTLE, PURCHASED BY ORGANIZATIONS, TOTAL, ANIMALS
9	CATTLE, PURCHASED FROM OUTSIDE THE REGION, ANIMALS
10	CATTLE, PURCHASED FROM OUTSIDE THE REGION, %
11	PURCHASED CATTLE, INCL. IMPORT, ANIMALS
12	PURCHASED CATTLE, INCL. IMPORT, %
13	DAIRY CATTLE, PURCHASED BY ORGANIZATIONS, TOTAL, ANIMALS
14	DAIRY CATTLE, PURCHASED FROM OUTSIDE THE REGION, ANIMALS
15	DAIRY CATTLE, PURCHASED FROM OUTSIDE THE REGION, %
16	PURCHASED DAIRY CATTLE, INCL. IMPORT, ANIMALS
17	PURCHASED DAIRY CATTLE, INCL. IMPORT, %

3.2. Подготовка исходных данных, формализация предметной области

Исходные данные для разработки интеллектуального облачного Эйдос-приложения взяты из официальной статистики региональных ветеринарных служб и имеют вид, приведенный в таблице 3:

Таблица 3. Исходные данные (фрагмент)

Years	The total number of patients: - Zareist.boln.KRS, including young (goals)	Total number of cases: -Forture. Slaughtered cattle, incl. young (goals)	Including: - diseases of the digestive system; -Registered. Sick cattle, including young (goals)	Including: - diseases of the digestive system; -vynuzhdenie. Slaughtered cattle, including young (goals)	- diseases of the respiratory system; -Zareistr.boln.KRS, including young (goals)	***	Acquired cattle from outside the region, goals	Acquired cattle from outside the region, %	Acquired Dairy, including Import%
2011 year	306700	3163	71725	362	4713	***	6715	91,7	81,6
2012 year	287633	4000	16837	427	5763	***	1829	67,4	66,3
2013 year	277828	3814	15276	388	6052	***	2427	89,8	74,1
2014 year	270790	0	11879	0	6959	***	2514	81,7	52,6
2015 year	267203	0	13259	0	11912	***	1739	88,4	0,5
2016 year	265216	2196	17634	347	8899	***	3102	81,2	4,8
2017 year	248683	3046	13050	819	7767	***	4418	67,8	36,1
2018 year	237729	2665	59022	907	34039	***	1140	39,1	0,0
2019 year	210801	2771	50173	974	27857	***	1466	58,6	23,0

Полный файл исходных данных имеет значительно большую размерность и не может быть приведен в данной работе из-за ограничения на ее объем. Но его можно скачать по ссылке из интеллектуального облачного Эйдос-приложения:

http://aidos.byethost5.com/Source_data_applications/Applications-000202/Inp_data.xls

Классификационные шкалы выделены желтым фоном.

Исходные данные содержат 3364 фактов. Каждый факт представляет собой наблюдение определенного значения фактора при переходе объекта моделирования в некоторое будущее состояние

3.3. Формализация предметной области

Для ввода этих исходных данных в систему «Эйдос» и формализации предметной области был использован автоматизированный программный интерфейс (API-2.3.2.2).

2.3.2.2. Universal software for data import interface in "EIDOS-X ++"

Automatic formalization domain: generating classification and descriptive scales and gradation, as well as training and recognize the sample on the basis of baseline data: "Inp_data"

Specify the file type of the original data: "Inp_data":

☒ XLS - MS Excel-2003 Standard XLS-file
☐ XLSX - MS Excel-2007 (2010) Standard DBF-file
☐ DBF - DBASE IV (DBF / NTX) CSV-file standard
☐ CSV - Comma-Separated Values

Set the parameters:

☒ Zeros and gaps is the lack of data
☐ Zeros and spaces count value of the data
☐ Create a database medium grade "Inp_davr.dbf"?

Requirements for raw data file

Specify the range of columns classification scales:

Initial column classification scales:
The final column of the classification scales:

Set the range of columns descriptive scales:

Basic descriptive column scales:
The final column of descriptive scales:

Set Mode:

☒ Formalization of the domain (on "Inp_data" basis)
☐ Generate detectable sample (for "Inp_rasp" basis)

Specify the method for selecting the size range:

☒ Regular intervals with a different number of observations
☐ Different intervals with an equal number of observations

the formation of a task scripting options or ways of interpreting the text fields "Inp_data":

☒ Do not use ACK scenario method analysis
☐ Apply spets.interpretatsiyu textboxes classes
☐ Apply scenario method ASA analysis
☐ Apply spets.interpretatsiyu text fields attributes

Parameters "Inp_data" interpretation of the text field values:

Interpretation TXT-class fields:

Field Values text file classification scales
raw data "Inp_data" are considered as a whole

Interpretation TXT-field attributes:

The values of the fields of descriptive text file scales
raw data "Inp_data" are considered as a whole

What are the names of the gradation numerical scales used:

☒ Only the interval numeric values (E.g. "1 / 3 - {59873.0000000, 178545.6666667}")
☐ Interval of numeric values (For example: "Minimum")
☐ And interval numeric values, and their names (E.g. "Minimum: 1 / 3 - {59873.0000000, 178545.6666667}")

Ok Cancel

Рисунок 1

Так как в данном приложении использованы числовые шкалы, то система запрашивает количество числовых диапазонов, на которое их разделять и мы задали 3 диапазона в классификационных шкалах и 5 в описательных шкалах. В результате в данном режиме автоматически разработаны градации классификационных и описательных шкал (таблицы 4 и 5), а затем с их использованием исходные данные закодированы в результате чего сформирована обучающая выборка (таблица 6):

Таблица 4. Классификационные шкалы и градации (фрагмент)

Код	Наименование
1	TOTAL NUMBER OF CASES: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)-1/3-{210801.0000000, 242767.3333333}
2	TOTAL NUMBER OF CASES: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)-2/3-{242767.3333333, 274733.6666667}
3	TOTAL NUMBER OF CASES: - REGISTERED SICK CATTLE, INCL. YOUNG CATTLE (ANIMALS)-3/3-{274733.6666667, 306700.0000000}
4	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)-1/3-{6248.0000000, 6636.6666667}
5	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)-2/3-{6636.6666667, 7025.3333333}
6	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (ANIMALS)-3/3-{7025.3333333, 7414.0000000}
7	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (%) -1/3-{1.6000000, 1.9333333}
8	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (%) -2/3-{1.9333333, 2.2666667}
9	TOTAL NUMBER OF CASES: - DIED CATTLE, INCL. YOUNG CATTLE (%) -3/3-{2.2666667, 2.6000000}
10	TOTAL NUMBER OF CASES: -COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)-1/3-{2196.0000000, 2797.3333333}
11	TOTAL NUMBER OF CASES: -COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)-2/3-{2797.3333333, 3398.6666667}
12	TOTAL NUMBER OF CASES: -COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (ANIMALS)-3/3-{3398.6666667, 4000.0000000}
13	TOTAL NUMBER OF CASES: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%) -1/3-{0.7200000, 0.9466667}
14	TOTAL NUMBER OF CASES: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%) -2/3-{0.9466667, 1.1733333}
15	TOTAL NUMBER OF CASES: - COMPULSORY SLAUGHTERED CATTLE, INCL. YOUNG CATTLE (%) -3/3-{1.1733333, 1.4000000}

Таблица 5. Описательные шкалы и градации (фрагмент)

Код	Наименование
1	HUNGARY-1/5-{136.0000000, 158.0000000}
2	HUNGARY-2/5-{158.0000000, 180.0000000}
3	HUNGARY-3/5-{180.0000000, 202.0000000}
4	HUNGARY-4/5-{202.0000000, 224.0000000}
5	HUNGARY-5/5-{224.0000000, 246.0000000}
6	AUSTRIA-1/5-{55.0000000, 155.4000000}
7	AUSTRIA-2/5-{155.4000000, 255.8000000}
8	AUSTRIA-3/5-{255.8000000, 356.2000000}
9	AUSTRIA-4/5-{356.2000000, 456.6000000}
10	AUSTRIA-5/5-{456.6000000, 557.0000000}
11	USA-1/5-{1409.0000000, 1561.2000000}
12	USA-2/5-{1561.2000000, 1713.4000000}
13	USA-3/5-{1713.4000000, 1865.6000000}
14	USA-4/5-{1865.6000000, 2017.8000000}
15	USA-5/5-{2017.8000000, 2170.0000000}
16	NETHERLANDS-1/5-{9.0000000, 139.2000000}
17	NETHERLANDS-2/5-{139.2000000, 269.4000000}
18	NETHERLANDS-3/5-{269.4000000, 399.6000000}
19	NETHERLANDS-4/5-{399.6000000, 529.8000000}
20	NETHERLANDS-5/5-{529.8000000, 660.0000000}

Полностью справочники классификационных и описательных шкал и градаций не приводятся, т.к. в первом из них 129 строк, а во втором - 85 строк.

Обучающая выборка по сути представляет собой нормализованную с помощью классификационных и описательных шкал и градаций (таблицы 4 и 5) базу исходных данных (таблица 3) и имеет такой же вид, только в ней вместо чисел и текстов приведены их коды.

В результате автоматической формализации предметной области подготовлены все необходимые условия для выполнения следующего этапа АСК-анализа: синтеза и верификации моделей.

3.4. Синтез и верификация моделей

Далее запускаем режим синтеза и верификации моделей 3.5. Создаваемая модель по размерности составляет около 1.2% от теоретически максимальной модели, которую можно создать и обработать в системе «Эйдос». Процесс синтеза и верификации моделей занял около часа.

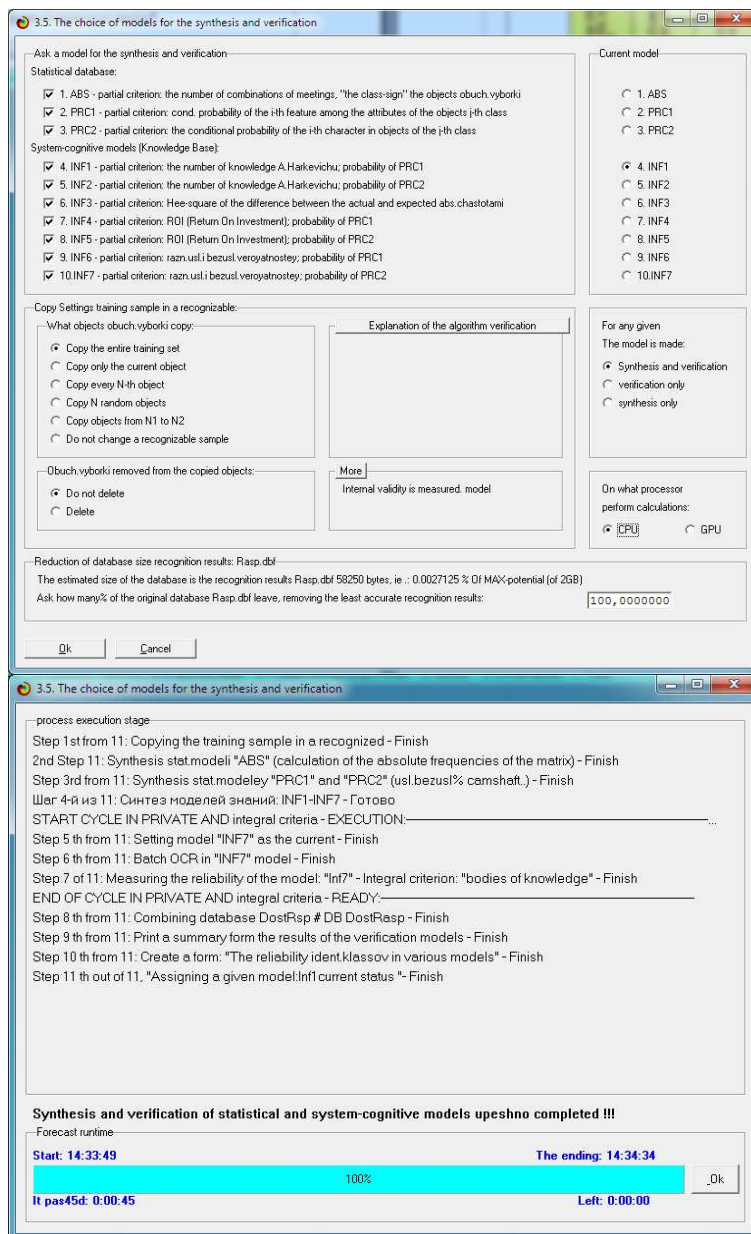


Рисунок 2

Для оценки достоверности моделей в системе «Эйдос» используется F-мера Ван Ризбергена и две ее улучшенные модификации, предложенные проф.Е.В.Луценко. По критерию L2 наилучшей по достоверности системно-когнитивной моделью является модель INF3 с интегральным критерием «Сумма знаний»: L2=0,926 при максимуме 1 (рисунок).

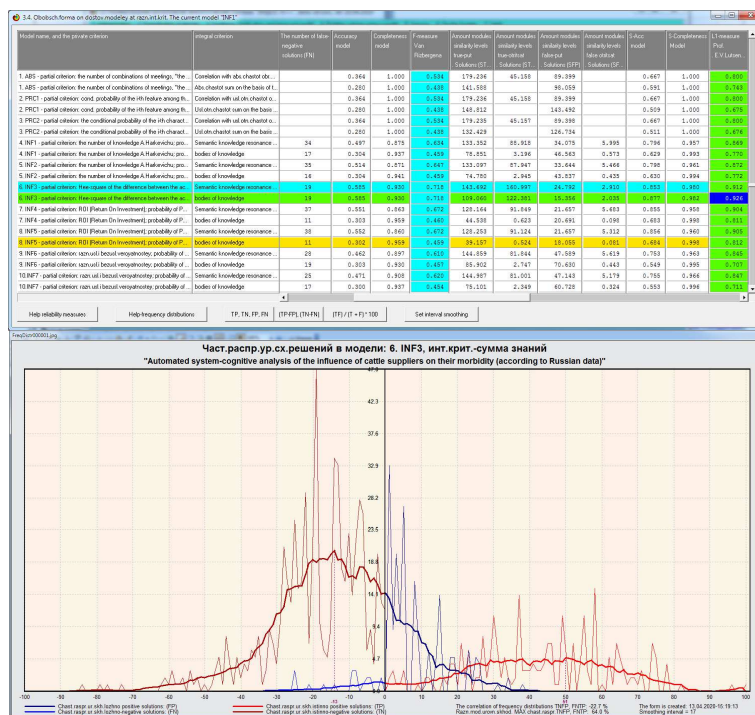


Рисунок 3

Это очень хороший результат, означающий, что созданные модели могут быть успешно применены для решения поставленных в работе задач. Поэтому данная модель и выбрана в качестве текущей для решения поставленных в работе задач на последующих этапах АСК-анализа.

3.5. Придание статуса текущей наиболее достоверной модели INF3

Далее в режиме 5.6 присвоим наиболее достоверной модели INF3 статус текущей модели и будем решать в ней все ранее поставленные задачи.

3.6. Решение задачи системной идентификации

При решении задачи идентификации для каждого объекта распознаваемой выборки в наиболее достоверной модели INF3 рассчитываются значения интегрального критериев для каждого класса. При этом определяется степень сходства каждого объекта с обобщенными образами всех классов, а потом для каждого объекта все классы ранжируются в порядке убывания сходства с объектом и, таким образом, прогнозируются объемы и структура заболеваемости крупного рогатого скота.

В системе «Эйдос» есть выходная форма (всего же их сотни), в которой приведены классы, ранжированные в порядке убывания сходства с выбранным текстом по двум интегральным критериям. Есть также форма в которой мы наоборот, видим тексты, ранжированные по степени сходства с заданными классами.

3.7. Решение задачи принятия решений (вывод информации о результатах многопараметрической типизации)

На этапе синтеза моделей путем обобщения примеров обучающей выборки были созданы обобщенные образы классов. Для научного исследования и принятия решений представляет интерес, а что же это за образы классов, какое их содержание. Можно вывести информацию об этом в форме SWOT-диаграмм, из которых видно какие значения факторов наиболее характерны (слева) и наиболее нехарактерны (справа) для данного класса, т.е. объема и вида заболевания.

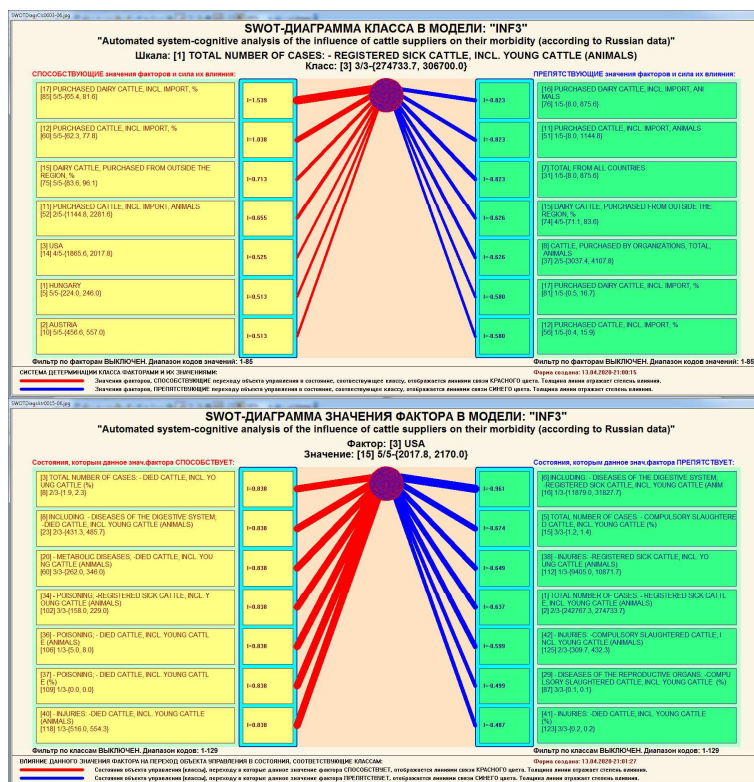


Рисунок 4

3.8. Решение задачи исследования моделируемой предметной области путем исследования ее модели

Если модель верно отражает моделируемую предметную область, то исследование модели обоснованно можно считать исследованием самой моделируемой предметной области.

В системе «Эйдос» исследование моделируемой предметной области включает: инвертированные SWOT-диаграммы значений факторов (семантические потенциалы слов); кластерно-конструктивный анализ классов; кластерно-конструктивный анализ значений факторов (слов); нелокальные нейроны; нелокальная нейронная сеть; 3D-интегральные когнитивные карты; 2D-интегральные когнитивные карты содержательного сравнения классов; 2D-интегральные когнитивные карты содержательного сравнения значений факторов (слов); когнитивные функции; значимость факторов и их значений; степень детерминированности классов и классификационных шкал.

В данной работе из-за ограничений на ее объем мы не будем рассматривать задачи исследования моделируемой предметной области путем исследования ее модели.

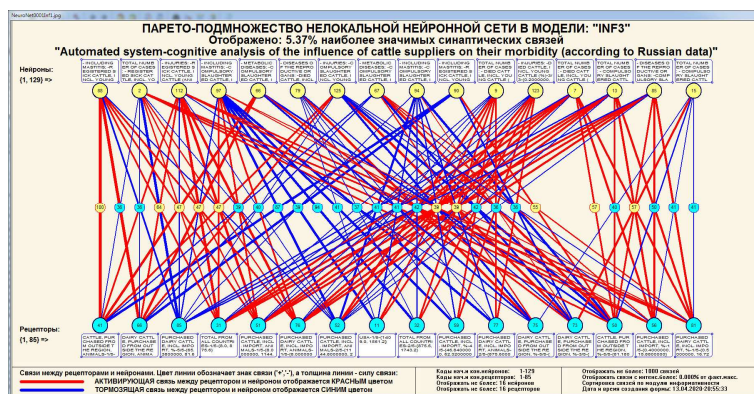


Рисунок 5

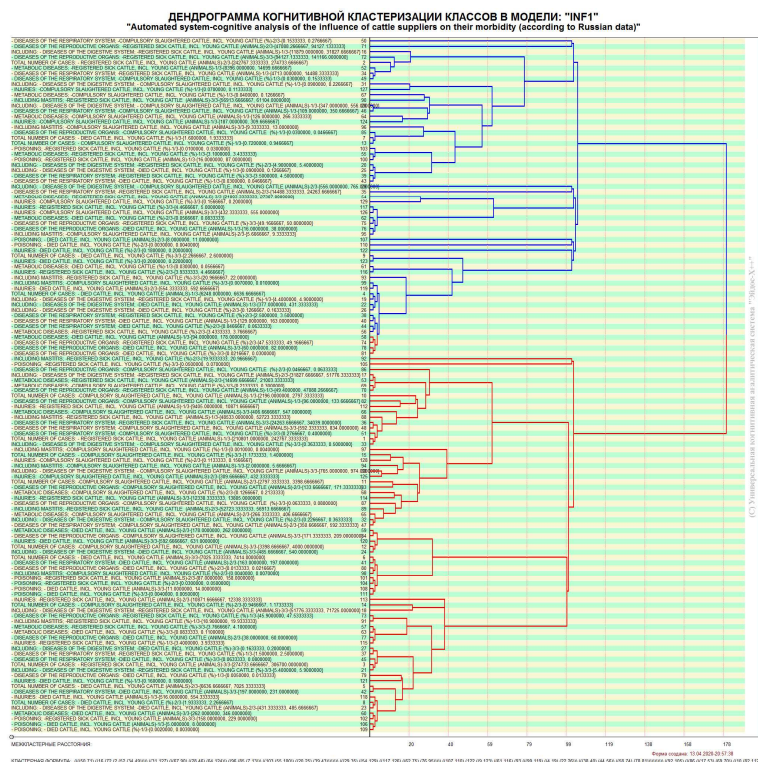


Рисунок 6

4. Обсуждение (Discussion)

По мнению авторов получены новые и интересные в научном плане результаты, имеющие практическое значение для принятия решений о выборе мест приобретения крупного рогатого скота. Эти результаты были получены путем применения современных интеллектуальных технологий, а конкретно – автоматизированного системно-когнитивного анализа и его программного инструментария – интеллектуальной системы «Эйдос».

По-видимому, полученные результаты надо признать предварительными, т.к. исходные данные имели сравнительно небольшой лонгитюд (9 лет) и были значительно фрагментированы. В дальнейшем при продолжении данного направления исследований и разработок рекомендуется обратить особое внимание на получение более полных данных.

Тем ни менее полученные результаты являются вполне достаточными для получения общей картины закономерностей в данной предметной области, т.е. они адекватно количественно отражают ветеринарные последствия решений о приобретении того или иного количества крупного рогатого скота у поставщиков из других регионов и стран. Сравнить полученные результаты с работами других ученых не представляется возможным, т.к. работы в этой области отсутствуют, тем более основанные на российских данных и для кубанского региона.

Автоматизированный системно-когнитивный анализ и его инструментарий позволяют корректно обрабатывать большие объемы фрагментированных и зашумленных данных, представленных в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения. Однако временной период, который охватывают эти данные необходимо увеличить. Для повышения научной и практической значимости исследования целесообразно провести подобное исследование в других регионах России. Возможно в созданных при этом моделях будет учтена региональная специфика и они будут в этих регионах работать лучше, чем модели, созданные на данных Краснодарского края. Вместе с тем в Краснодарском крае можно пользоваться уже полученными результатами.

При написании данной статьи были изучены многочисленные работы различных авторов в области ветеринарии. Однако ни в одной из этих статей или работ не приводится математическая модель, предложенная в данной статье и не используются предложенные в ней частные и интегральные критерии (1)-(6). Поэтому данная статья представляет интерес для исследователей в данной предметной области и является актуальной.

5. Выводы (Conclusions)

5.1. Основные результаты

Данная работа демонстрирует, что математические модели (частные и интегральные критерии), методики численных расчетов (структуры данных и алгоритмы их обработки), экранные формы управления процессами, программный интерфейс ввода исходных данных в систему (API) «Эйдос» и повышения степени формализации исходных данных до нормализованных баз данных, экранные формы текстовых и графических выходных форм по результатам решения задач прогнозирования, принятия решений и исследования, программная реализация математических моделей, методик численных расчетов, интерфейса и когнитивной графики в системе «Эйдос» являются адекватным средством решения поставленных в работе проблемы и задач.

5.2. Итог статьи

Как показывает анализ результатов численного эксперимента предложенное и реализованное в системе «Эйдос» решение поставленных задач является эффективным, что позволяет обоснованно утверждать, что цель работы достигнута. Любой желающий может лично ознакомиться с интеллектуальным Эйдос-приложением, т.к. оно размещено в Эйдос-облаке под номером 202:

http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.htm

Сама система «Эйдос» относится к бесплатному открытому программному обеспечению. Ее можно скачать по ссылке: <http://lc.kubagro.ru/aidos/Aidos-X.htm>. Установить данное приложение можно в Диспетчере приложений (режим 1.3).

Список литературы (References)

1. 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.
2. Луценко Е.В. Автоматизированный системно-когнитивный анализ в ветеринарии (на примере разработки диагностических тестов) / Е.В. Луценко // Политематический 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, doi: [10.21515/1990-4665-137-031](https://doi.org/10.21515/1990-4665-137-031). – Режим доступа: <http://ej.kubagro.ru/2018/03/pdf/31.pdf>, 3,375 у.п.л.
3. Луценко Е.В. Агломеративная когнитивная кластеризация нозологических образов 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, doi: [10.21515/1990-4665-138-033](https://doi.org/10.21515/1990-4665-138-033). – Режим доступа: <http://ej.kubagro.ru/2018/04/pdf/33.pdf>, 1,125 у.п.л.
4. Луценко Е.В. Агломеративная когнитивная кластеризация симптомов и синдромов в ветеринарии / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №05(139). С. 99 – 116. – IDA [article ID]: 1391805033, doi: [10.21515/1990-4665-139-033](https://doi.org/10.21515/1990-4665-139-033). – Режим доступа: <http://ej.kubagro.ru/2018/05/pdf/33.pdf>, 1,125 у.п.л.
5. Луценко Е.В. Автоматизированный системно-когнитивный анализ антибиотиков в ветеринарии / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №06(140). С. 171 – 220. – IDA [article ID]: 1401806033, doi: [10.21515/1990-4665-140-033](https://doi.org/10.21515/1990-4665-140-033). – Режим доступа: <http://ej.kubagro.ru/2018/06/pdf/33.pdf>, 3,125 у.п.л.
6. Луценко Е.В. Разработка ветеринарного теста для диагностики желудочно-кишечных заболеваний лошади на основе данных репозитория UCI с применением АСК-анализа / Е.В. Луценко, Е.К. Печурина, А.Э. Сергеев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №07(141). С. 111 – 175. – IDA [article ID]: 1411807033, doi: [10.21515/1990-4665-141-033](https://doi.org/10.21515/1990-4665-141-033). – Режим доступа: <http://ej.kubagro.ru/2018/07/pdf/33.pdf>, 4,062 у.п.л.
7. Луценко Е.В. Автоматизированный системно-когнитивный анализ и классификация пород крупного рогатого скота / Е.В. Луценко, Е.К. Печурина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №08(142). С. 68 – 95. – IDA [article ID]: 1421808033, doi: [10.21515/1990-4665-142-033](https://doi.org/10.21515/1990-4665-142-033). – Режим доступа: <http://ej.kubagro.ru/2018/08/pdf/33.pdf>, 1,75 у.п.л.

8. Луценко Е.В. Математическое и численное моделирование взаимосвязи морфологического, биохимического и микроэлементного состава крови бычков герефордской породы и их размеров / Е.В. Луценко, В.Г. Лежнев, Н.И. Ковелин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №09(143). С. 49 – 88. – IDA [article ID]: 1431809033, doi: [10.21515/1990-4665-143-033](https://doi.org/10.21515/1990-4665-143-033). – Режим доступа: <http://ej.kubagro.ru/2018/09/pdf/33.pdf>, 2,5 у.п.л.

9. Луценко Е.В. Формирование семантического ядра ветеринарии путем Автоматизированного системно-когнитивного анализа паспортов научных специальностей ВАК РФ и автоматическая классификация текстов по направлениям науки / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №10(144). С. 44 – 102. – IDA [article ID]: 1441810033, doi: [10.21515/1990-4665-144-033](https://doi.org/10.21515/1990-4665-144-033). – Режим доступа: <http://ej.kubagro.ru/2018/10/pdf/33.pdf>, 3,688 у.п.л.

10. Луценко Е.В. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификация статей по научным специальностям с применением АСК-анализа и интеллектуальной системы «Эйдос» (на примере Научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии) / Е.В. Луценко, Н.В. Андрафанова, Н.В. Потапова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2019. – №01(145). С. 31 – 102. – IDA [article ID]: 1451901033, doi: [10.21515/1990-4665-145-033](https://doi.org/10.21515/1990-4665-145-033). – Режим доступа: <http://ej.kubagro.ru/2019/01/pdf/33.pdf>, 4,5 у.п.л.

11. Луценко Е.В. Когнитивная ветеринария – ветеринария цифрового общества: дефиниция базовых понятий / Е.В. Луценко, Е.К. Печурина, А.Э. Сергеев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2019. – №08(152). С. 141 – 199. – IDA [article ID]: 1521908015. – Режим доступа: <http://ej.kubagro.ru/2019/08/pdf/15.pdf>, 3,688 у.п.л.