

УДК 004.8

5.2.2. Математические, статистические и инструментальные методы экономики (физико-математические науки, экономические науки)

EDA, KDD И СИСТЕМНО-КОГНИТИВНЫЕ НЕЙРОСЕТИ – НОВЫЕ ПЕРСПЕКТИВНЫЕ МОДЕЛИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЭКОНОМИКЕ

Луценко Евгений Вениаминович
д.э.н., к.т.н., профессор
Web of Science ResearcherID S-8667-2018
Scopus Author ID: 57188763047
РИНЦ SPIN-код: 9523-7101

prof.lutsenko@gmail.com <http://lc.kubagro.ru>
https://www.researchgate.net/profile/Eugene_Lutsenko

*Кубанский Государственный Аграрный университет
имени И.Т.Трубилина, Краснодар, Россия*

В современной экономике все больше внимания уделяется применению математических и статистических методов. Разрабатываются инструментальные методы, т.е. программное обеспечение, реализующее эти методы. В старой номенклатуре научных специальностей ВАК РФ была специальность: 08.00.13 - Математические и инструментальные методы экономики. В новой номенклатуре научных специальностей ВАК РФ эта специальность называется: 5.2.2. Математические, статистические и инструментальные методы в экономике. В условиях современной революции в искусственном интеллекте актуальной становится новая специальность: 5.2.8. Информационные, когнитивные и интеллектуальные технологии в экономике. В рамках этой специальности разрабатываются технологии преобразования данных в информацию, а ее в знания и применения этих знаний для решения различных экономических задач. Основные из этих задач это идентификация и прогнозирование состояний экономических систем, поддержка принятия экономических решений, исследование экономического объекта моделирования и управления путем исследования его интеллектуальной модели. Для этих целей используется много различных типов систем. 1. Системы управления базами данных (СУБД): Эти системы позволяют организовывать, хранить и извлекать данные для последующей обработки и анализа. 2. Бизнес-аналитические системы: Они используются для анализа данных и выявления ключевых трендов и паттернов, которые могут быть преобразованы в информацию и знания для принятия решений. 3. Системы искусственного интеллекта и машинного обучения: Эти системы могут

UDC 004.8 UDC 004.8

5.2.2. Mathematical, statistical and instrumental methods of economics (physical and mathematical sciences, economic sciences)

EDA, KDD AND SYSTEM-COGNITIVE NEURAL NETWORKS ARE NEW PROMISING MODELS OF ARTIFICIAL INTELLIGENCE IN ECONOMICS

Lutsenko Evgeniy Veniaminovich
Doctor of Economics, Cand.Tech.Sci., Professor
Web of Science ResearcherID S-8667-2018
Scopus Author ID: 57188763047
RSCI SPIN-code: 9523-7101

prof.lutsenko@gmail.com <http://lc.kubagro.ru>
https://www.researchgate.net/profile/Eugene_Lutsenko

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

In modern economics, more and more attention is being paid to the application of mathematical and statistical methods. Instrumental methods are being developed, i.e. software implementing these methods. In the old nomenclature of scientific specialties of the Higher Attestation Commission of the Russian Federation there was a specialty: 08.00.13 - Mathematical and instrumental methods of economics. In the new nomenclature of scientific specialties of the Higher Attestation Commission of the Russian Federation, this specialty is called: 5.2.2. Mathematical, statistical and instrumental methods in economics. In the context of the modern revolution in artificial intelligence, a new specialty is becoming relevant: 5.2.8. Information, cognitive and intellectual technologies in economics. Within the framework of this specialty, technologies are being developed for converting data into information, and its into knowledge and applying this knowledge to solve various economic problems. The main of these tasks are the identification and forecasting of the states of economic systems, support for economic decision-making, and the study of an economic object of modeling and management through the study of its intellectual model. Many different types of systems are used for these purposes. 1. Database Management Systems (DBMS): These systems allow you to organize, store and retrieve data for subsequent processing and analysis. 2. Business analytical systems: They are used to analyze data and identify key trends and patterns that can be transformed into information and knowledge for decision-making. 3. Artificial intelligence and machine learning systems: These systems can automatically analyze data, identify patterns and make predictions based on available information, which allows you to gain new knowledge and improve decision-making processes.

автоматически анализировать данные, выявлять закономерности и делать прогнозы на основе имеющейся информации, что позволяет получить новые знания и улучшить процессы принятия решений. Но особое место среди систем преобразования данных в информацию, а ее в зания занимают системы EDA (Exploratory Data Analysis) и KDD (Knowledge Discovery in Databases). Эти системы непрерывно совершенствуются во всем мире. Предлагается новое перспективное направлений совершенствования этих систем, основанное на применении в качестве исходных данных для них не эмпирических данных, а системно-когнитивных моделей автоматизированного системно-когнитивного анализа (АСК-анализ), генерируемых его программным инструментарием - интеллектуальной системой "Эйдос"

Ключевые слова: EDA, KDD, СК-МОДЕЛИ, АСК-АНАЛИЗ, АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА «ЭЙДОС»

But EDA (Exploratory Data Analysis) and KDD (Knowledge Discovery in Databases) systems occupy a special place among data-to-information conversion systems, and its applications. These systems are continuously being improved all over the world. A new promising direction for improving these systems is proposed, based on the use of system-cognitive models of automated system-cognitive analysis (ASK-analysis) generated by its software tools - the intelligent Eidos system, not empirical data, as initial data for them

Keywords: EDA, KDD, SK MODELS, ASK ANALYSIS, AUTOMATED SYSTEM-COGNITIVE ANALYSIS, INTELLIGENT SYSTEM "EIDOS"

<http://dx.doi.org/10.21515/1990-4665-198-017>

Введение

В современном мире экономических исследований и практики все большее внимание уделяется применению методов искусственного интеллекта (ИИ) для анализа данных и выявления в них причинно-следственных закономерностей. Среди таких методов особое место занимают системы исследования данных (EDA), процесс открытия знаний в базах данных (KDD) и системно-когнитивные нейросети. В этой работе мы рассмотрим подробные аспекты каждого из этих методов и их применение в экономике, а также обсудим преимущества использования системно-когнитивных моделей для повышения достоверности моделей знаний, созданных в результате анализа данных.

Данные, информация и знания

Данные, информация и знания - это три понятия, тесно связанные друг с другом, но имеющие различные уровни значимости и обработки:

<http://ej.kubagro.ru/2024/04/pdf/17.pdf>

1. Данные: Это фактические фрагменты информации, либо набор символов или цифр без контекста или интерпретации. Данные могут представлять собой простой поток информации, который еще не обработан или организован для использования.

2. Информация: Это данные, которые уже обработаны и организованы таким образом, чтобы иметь смысл или ценность для получателя. Информация содержит контекст и имеет возможность передавать определенное значение или понимание.

3. Знания: Это высший уровень абстракции, получаемый из информации путем анализа, интерпретации и выводов. Знания представляют собой не только понимание определенной информации, но и умение применять ее в различных контекстах и сделать выводы.

В экономике преобразование данных в информацию и знания играет решающую роль в принятии решений, анализе трендов, прогнозировании и планировании. Например:

- Преобразование данных о рыночных ценах в информацию об изменениях цен на товары и услуги позволяет компаниям анализировать конкурентоспособность своих продуктов.

- Анализ финансовых данных позволяет выявлять тенденции в расходах и доходах, что помогает компаниям принимать решения о финансовой стратегии и бюджетировании.

- Преобразование информации о потребительских предпочтениях в знания о том, что именно заставляет клиентов предпочитать одни товары другим, помогает разрабатывать маркетинговые стратегии и улучшать продукты.

Для преобразования данных в информацию и знания в экономике используются различные типы систем, включая:

1. Системы управления базами данных (СУБД): Эти системы позволяют организовывать, хранить и извлекать данные для последующей обработки и анализа.

2. Бизнес-аналитические системы: Они используются для анализа данных и выявления ключевых трендов и паттернов, которые могут быть преобразованы в информацию и знания для принятия решений.

3. Системы искусственного интеллекта и машинного обучения: Эти системы могут автоматически анализировать данные, выявлять закономерности и делать прогнозы на основе имеющейся информации, что позволяет получить новые знания и улучшить процессы принятия решений.

Но особое место среди систем преобразования данных в информацию, а ее в зания занимают системы EDA (Exploratory Data Analysis) и KDD (Knowledge Discovery in Databases). EDA и KDD – это два ключевых подхода к анализу данных, которые используются для извлечения ценной информации и знаний из наборов данных. Вот их краткое описание.

EDA (Exploratory Data Analysis):

- EDA представляет собой метод исследования данных с целью понимания их структуры, свойств и характеристик.

- Цель EDA - выявить ключевые особенности данных, выявить связи и закономерности между переменными, обнаружить аномалии и выбросы, а также определить наиболее подходящие методы анализа.

- Методы EDA включают в себя визуализацию данных, описательную статистику, корреляционный анализ, анализ распределения и многое другое.

- EDA помогает исследователям получить интуитивное понимание данных до применения более сложных аналитических методов.

KDD (Knowledge Discovery in Databases):

- KDD представляет собой процесс извлечения ценной информации и знаний из больших объемов данных.

- Этот процесс включает в себя несколько этапов, таких как отбор и предварительная обработка данных, преобразование и обогащение данных, поиск шаблонов и закономерностей, оценка и интерпретация результатов.

- Цель KDD - не только найти интересные закономерности в данных, но и создать модели и инструменты, которые могут использоваться для прогнозирования, классификации, кластеризации и других задач анализа данных.

- В ходе процесса KDD могут использоваться различные методы и технологии, включая машинное обучение, статистический анализ, интеллектуальный анализ данных и многое другое.

В целом, EDA и KDD представляют собой ключевые шаги в процессе анализа данных, начиная от первичного исследования до извлечения ценных знаний и информации, которые могут быть использованы для принятия решений и развития бизнеса.

Применение этих систем позволяет компаниям эффективно использовать данные, преобразовывая их в информацию и знания, что в свою очередь способствует более точному анализу ситуации на рынке, принятию обоснованных решений и сокращению рисков.

Рассмотрим EDA и KDD системы подробнее.

1. Системы исследования данных (EDA)

Системы Исследования Данных (EDA)

Системы исследования данных (EDA) представляют собой процесс анализа данных с целью выявления закономерностей, связей и тенденций. EDA является важным этапом в процессе анализа данных, предшествующим

построению моделей и принятию решений. Вот более подробное описание этого процесса:

1. Визуализация данных:

Визуализация данных является одним из ключевых этапов EDA. На этом этапе данные преобразуются в графические формы для визуального анализа. Это может включать построение различных типов графиков, таких как гистограммы, диаграммы рассеяния, ящики с усами и тепловые карты. Визуализация данных позволяет исследователям быстро обнаружить общие распределения данных, выбросы, корреляции и другие интересные закономерности.

2. Статистический анализ:

На этом этапе данные анализируются с использованием различных статистических методов. Это может включать вычисление описательных статистик, таких как среднее значение, медиана, стандартное отклонение и квартили. Кроме того, проводится корреляционный анализ для определения взаимосвязей между переменными. Статистический анализ позволяет исследователям получить количественное представление о данных и выявить значимые характеристики.

3. Агрегация данных:

На этом этапе данные агрегируются или суммируются для создания обобщенных представлений. Например, данные могут быть агрегированы по временным интервалам, регионам или категориям, чтобы получить общие тренды и паттерны. Агрегация данных помогает упростить анализ и выделить основные характеристики набора данных.

4. Поиск аномалий и выбросов:

На этом этапе данные анализируются на наличие аномалий или выбросов, которые могут исказить результаты анализа. Это может включать

выявление необычных значений, отклоняющихся от общего распределения данных. Идентификация аномалий позволяет исследователям понять причины этих отклонений и принять меры по их исправлению.

5. Интерактивный анализ:

Это последний этап EDA, который включает интерактивный анализ данных с использованием специализированных инструментов и приложений. Интерактивный анализ позволяет исследователям проводить более глубокое и динамичное исследование данных, взаимодействуя с ними в режиме реального времени. Это может включать использование интерактивных диаграмм, фильтров и инструментов для исследования конкретных аспектов данных.

EDA играет важную роль в процессе анализа данных, позволяя исследователям получить предварительное представление о данных, выявить важные характеристики и обнаружить потенциальные проблемы или аномалии. Хорошо выполненный анализ данных на этом этапе может значительно улучшить качество последующих шагов анализа и принятия решений.

2. Процесс открытия знаний в базах данных (KDD)

Системы преобразования данных в знания, также известные как процесс открытия знаний в базах данных (Knowledge Discovery in Databases - KDD), представляют собой циклический процесс, включающий несколько этапов для автоматизированного извлечения полезной информации из больших объемов данных и преобразования ее в знания, которые могут быть использованы для принятия решений и построения моделей. Давайте рассмотрим подробнее каждый этап этого процесса:

1. Выборка данных:

Этот этап включает выборку данных из источников данных для анализа. Выборка должна быть репрезентативной и включать все необходимые переменные для достижения целей исследования. При выборе данных необходимо учитывать их качество, актуальность и доступность.

2. Предварительная обработка данных:

На этом этапе данные проходят через ряд преобразований, включая очистку, трансформацию и редукцию размерности. Очистка данных включает в себя удаление дубликатов, заполнение пропущенных значений и коррекцию ошибок. Трансформация данных может включать стандартизацию, нормализацию и преобразование категориальных переменных. Редукция размерности помогает уменьшить количество признаков, сохраняя при этом значимую информацию.

3. Извлечение признаков:

На этом этапе выделяются наиболее важные характеристики данных для построения моделей. Это может включать выделение главных компонент, отбор признаков и создание новых признаков на основе существующих. Извлечение признаков позволяет сократить размерность данных и выявить скрытые закономерности.

4. Моделирование:

На этом этапе применяются алгоритмы машинного обучения для построения прогностических моделей на основе данных. Это может включать классификацию, регрессию, кластеризацию и другие методы анализа данных. Моделирование позволяет выявить закономерности в данных и предсказать будущие значения.

3. Системно-когнитивные нейросети: нелокальные сети прямого счета

Системно-когнитивные нейросети предложены автором в работе [1] представляют собой современный подход к разработке нейронных сетей, который учитывает принципы системного и когнитивного моделирования. В отличие от традиционных нейронных сетей, которые обычно ориентированы на обработку статистических данных, системно-когнитивные нейросети стремятся моделировать сложные взаимодействия между элементами системы и учитывать психологические и когнитивные аспекты принятия решений.

Архитектура системно-когнитивных нейросетей: Эти нейросети обычно имеют более сложную архитектуру, чем традиционные нейронные сети. Они могут включать различные слои, в том числе слои, имитирующие когнитивные процессы, такие как внимание, память и принятие решений.

Обучение системно-когнитивных нейросетей: Для обучения таких нейросетей часто применяются методы обучения с подкреплением и обучения с учителем. Это позволяет моделировать сложные взаимодействия между элементами системы и учитывать различные аспекты принятия решений.

Применение в экономике: В экономике системно-когнитивные нейросети могут использоваться для моделирования поведения экономических агентов, прогнозирования рыночных тенденций и оптимизации экономических стратегий. Они позволяют учитывать не только статистические закономерности, но и психологические, социальные и когнитивные аспекты принятия решений.

5. Оценка и интерпретация результатов:

На последнем этапе происходит оценка качества моделей и интерпретация результатов их работы. Это включает анализ точности

моделей, проверку их на новых данных и интерпретацию полученных знаний. Оценка результатов позволяет убедиться в правильности принятых решений и определить дальнейшие шаги исследования.

Кроме того, важным аспектом процесса KDD является его итеративный характер. После оценки и интерпретации результатов исследователи могут вернуться на предыдущие этапы для уточнения моделей и повышения их качества. Это позволяет создавать более точные и надежные модели на основе данных.

4. Преимущества системно-когнитивных моделей EDA и KDD

Традиционно в анализе данных используются эмпирические данные, которые могут быть неполными или содержать шум. Вместо этого, предлагается использовать системно-когнитивные модели, которые учитывают не только статистические закономерности, но и психологические, социальные и когнитивные аспекты поведения экономических агентов.

Системно-когнитивные модели могут помочь улучшить достоверность прогнозов и принятие решений в экономике, так как они способны учитывать более широкий спектр факторов и взаимосвязей между ними. Это позволяет получать более точные и надежные результаты, что особенно важно в условиях неопределенности и изменчивости современного мира.

Особо необходимо отметить быструю сходимость процессов обучения СК-нейросетей и вытекающую из этой их особенности высокую эффективность на малых выборках.

Заключение

Системы исследования данных (EDA), процесс открытия знаний в базах данных (KDD) и системно-когнитивные нейросети представляют собой мощные инструменты для анализа данных и разработки моделей искусственного интеллекта в экономике. Применение системно-когнитивных

моделей может значительно улучшить качество анализа данных и прогнозирования в экономических исследованиях и практике.

Литература

1. 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. – 2003. – № 1. – С. 76-88. – EDN JWXLKT.
2. 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. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.
3. Луценко, Е. В. Синтез адаптивных интеллектуальных измерительных систем с применением АСК-анализа и системы "Эйдос", системная идентификация в эконометрике, биометрии, экологии, педагогике, психологии и медицине / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2016. – № 116. – С. 1-60. – EDN VQUVHJ.
4. Луценко, Е. В. Революция начала XXI века в искусственном интеллекте: глубинные механизмы и перспективы / Е. В. Луценко, Н. С. Головин. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2024. – 394 с. – DOI 10.13140/RG.2.2.17056.56321. – EDN OMIPIL.

References

1. Lucenko, E. V. Sistemnaya teoriya informacii i nelokal'ny'e interpretiruemye nejronny'e seti pryamogo scheta / E. V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. – 2003. – № 1. – S. 76-88. – EDN JWXLKT.
3. Lucenko, E. V. Invariantnoe otnositel'no ob`emov danny`x nechetkoe mul'tiklassovoe obobshhenie F-mery` dostovernosti modelej Van Rizbergena v ASK-analize i sisteme "E`jdos" / E. V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. – 2017. – № 126. – S. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.
4. Lucenko, E. V. Sintez adaptivny`x intellektual'ny`x izmeritel'ny`x sistem s primeneniem ASK-analiza i sistemy` "E`jdos", sistemnaya identifikaciya v e`konometrike, biometrii, e`kologii, pedagogike, psixologii i medicine / E. V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. – 2016. – № 116. – S. 1-60. – EDN VQUVHJ.
6. Lucenko, E. V. Revolyuciya nachala XXI veka v iskusstvennom intellekte: glubinny`e mexanizmy` i perspektivy` / E. V. Lucenko, N. S. Golovin. – Krasnodar : Kubanskij gosudarstvenny`j agrarny`j universitet im. I.T. Trubilina, 2024. – 394 s. – DOI 10.13140/RG.2.2.17056.56321. – EDN OMIPIL.