**УДК 004.891**

**МЭС  ДЛЯ ВЫБОРА СИСТЕМ ПО РАБОТОТЕ С ОЗЁРАМИ ДАННЫХ И НЕЙРОСЕТЯМИ**

Ларионова А.П., Коноваликова С.А., Алёшин А.Д., Балабас А.Г.

ФГБОУ ВО МГТУ им. Н.Э. Баумана, г. Москва, Россия

***Аннотация.*** *В наши дни все вокруг говорят про пользу Big Data. В итоге бизнес пытается работать с масштабными базами данных, но сталкивается с проблемой — все данные разнородные и неструктурированные, перед загрузкой в базы их нужно долго обрабатывать. В итоге работа с Big Data оказывается слишком сложной и дорогой, а часть данных теряется, хотя могла бы принести пользу в будущем. Помочь с этим могут Data Lake — озера данных, которые помогают быстро и недорого работать с большими объемами неструктурированных данных. А добавление искусственного интеллекта может использоваться, чтобы делать прибыльные прогнозы. Наша работа посвящена выбору наиболее эффективной системы по работе с озёрами данных и выбору ПО для работы с нейросетями. Создана новая математическая модель данной предметной области, которая реализована в виде миварной экспертной системы (МЭС). Данная модель, реализована в виде миварной сети, включающей 31 правило и 60 параметров, эволюционно могут добавляться новые правила и параметры. Проект разработан для дальнейшей загрузки в среду КЭСМИ Wi!Mi Разуматор, в котором планируется развитие экспертной системы.*

***Ключевые слова:*** *мивар, миварные системы, экспертные системы, КЭСМИ, Wi!Mi, Разуматор, искусственный интеллект, Data Lake, Big Data, нейросеть, базы данных.*

**MES FOR SELECTING SYSTEMS FOR WORKING WITH DATA LAKES AND NUERAL NETWORKS**

Larionova A.P., Konovalikova S.A., Alyoshin A.D., Balabas A.G.

BMSTU, Moscow, Russia

***Abstract:*** *Nowadays, everyone around is talking about the benefits of Big Data. As a result, the business is trying to work with large-scale databases, but is faced with a problem - all data is heterogeneous and unstructured, and it takes a long time to process it before loading it into the database. As a result, working with Big Data turns out to be too complicated and expensive, and part of the data is lost, although it could be useful in the future. Data Lake can help with this - data lakes that help you quickly and inexpensively work with large amounts of unstructured data. And the addition of artificial intelligence can be used to make profitable predictions. Our work is devoted to choosing the most effective system for working with data lakes and choosing software for working with neural networks. A new mathematical model of this subject area has been created, which is implemented in the form of a mivar expert system (MES). This model is implemented as a mivar network that includes 31 rules and 60 parameters; new rules and parameters can be added evolutionarily. The project has been developed for further loading into the environment of Wi!Mi Razumator, in which the development of an expert system is planned.*

***Keywords:*** *mivar, mivar systems, expert systems, KESMI, Wi!Mi, Razumator, artificial intelligence, Data Lake, Big Data, neural network, databases.*

**Актуальность использования миварных систем.** Проблема создания интеллектуальных систем остается актуальной и практически значимой.

Миварные сети открывают принципиально новые возможности в сфере разработки действительно интеллектуальных и логически рассуждающих систем (ЛРС, Разуматоров) [2-9], а также позволяют перейти к новому поколению экспертных систем, автоматизированных систем управления (АСУ), систем поддержки принятия решений (СППР), распознавания образов, понимания, аннотирования и перевода текста, поисковых систем, при разработке роботов и др. Путем объединения баз данных с логико-вычислительной обработкой миварный подход позволяет работать с огромными объемами контекстной информации, что порождает способность компьютеров к непосредственному пониманию смысла информации и дает предпосылки создания логического искусственного интеллекта (ЛИИ). Миварный путь создания ИИ можно показать так [3]:

Миварные технологии -> БД + Логика -> Контекст -> Смысл -> Логический искусственный интеллект

Таким образом, миварные сети обусловливают переход к новому поколению экспертных систем и интеллектуальных пакетов прикладных программ [4]. Миварный подход позволил на практике создать автоматические, обучаемые, эволюционные, активные, логически рассуждающие информационные системы. В перспективе на основе миварных сетей будет создана глобальная мультипредметная активная экспертная система под названием «Миварная активная энциклопедия».

**Цели и задачи.** Необходимо оптимизировать этап планирования для жизненного цикла ПО; создать базу знаний для миварной экспертной системы разработки ПО для выбора варианта хранения информации для озера данных и выбора фреймворка для нейросети; рассмотреть и учесть факторы, влияющие на разработку программного обеспечения и проанализировать методологии, по которым разрабатывается программный продукт.

**Актуальность по ПО.** Миварный подход, в отличие от других научных исследований, может работать с глобальными познающе-диагностическими системами реального времени [2]. Для миварного подхода сразу было поставлено условие работы с открытыми и достаточно большими предметными областями, где счет логическим правилам идет на десятки тысяч.

Суть миварного подхода заключается в объединении баз данных и систем логико-вычислительной обработки в единые эволюционно развивающиеся системы, позволяющие собрать вместе различные научные разработки на основе сервисно-ориентированных архитектур и технологий интеллектуальных агентов (многоагентных систем). С появлением такого подхода становится реальным собрать воедино такие элементы, как базы данных, базы знаний, экспертные системы, многоагентные системы, сервисно-ориентированные архитектуры и др., и это позволит создать глобальные самоорганизующиеся программноаппаратные комплексы с эволюционной структурой для познающе-диагностических и экспертных систем, а также систем оперативной диагностики.

Поэтому миварные экспертные системы отлично подходят для работы с большими базами данных и нейросетями.

Для исследования озера данных, нам необходимо работать с различными видами представления данных, несколькими видами ПО и производить оценку скорости выполнения запросов, для этого мы будем использовать миварный подход, так как он предоставляет возможность хранить для одной предметной области несколько разных процедур, решающих одинаковые задачи. Более того, если вычислительные ресурсы позволяют, то все эти процедуры могут запускаться одновременно, а потом на основе конкуренции из представленных ими результатов выбирается наиболее подходящий для каждого конкретного случая.

Для работы с нейросетями также хорошо подойдет миварная система.

**База знаний**

|  |
| --- |
| Объекты |
| Графовая база данных |
| Реляционная база данных |
| Многомерная база данных |
| ПО Spark |
| ПО Neo4j |
| ПО Postgres |
| ПО Pentaho |
| нейросеть |
| аппаратная платформа Gyrfalcon |
| ПО Gyrfalcon |
| аппаратная платформа Khadas |
| ПО Khadas |

Таблица 1. Объекты базы знаний

Формализованные правила:

1. Если нужно реализовать и использовать графовую базу данных то нужно использовать специальное ПО: Spark и Neo4j
2. Если нужно реализовать и использовать реляционную базу данных, то нужно использовать специальное ПО: Spark и Postgres;

21. Если нужно реализовать и эффективно использовать нейросеть с использованием аппаратной платформы Gyrfalcon, то нужно использовать специальное ПО: Gyrfalcon;

22. Если нужно реализовать и эффективно использовать нейросеть с использованием аппаратной платформы Khadas, то нужно использовать специальное ПО: Khadas;

|  |  |  |  |
| --- | --- | --- | --- |
| Номер правила | Вход | Правило | Выход |
| 1 | графовая база данных | использовать | ПО Spark  ПО Neo4j |
| 2 | реляционная база данных | использовать | ПО Spark  ПО Postgres |
| 20 | нейросеть  аппаратная платформа Gyrfalcon | использовать | ПО Gyrfalcon |
| 21 | нейросеть  аппаратная платформа Khadas | использовать | ПО Khadas |

Таблица 2. Правила

**Заключение**. Была выбрана наилучшая методология разработки программных средств. Обозначены факторы, оказывающие влияние на создание программных продуктов. Создана база знаний для миварной экспертной системы, определяющая трудоемкость выполнения задач по разработке ПО, вероятность и оптимальное время их выполнения. Полученную базу знаний можно загружать в КЭСМИ.

**Список литературы**

1. Варламов О.О. Эволюционные базы данных и знаний для адаптивного синтеза интеллектуальных систем. Миварное информационное пространство. - М.: Радио и связь, 2002. - 288 c.
2. Варламов О.О. Системный анализ и синтез моделей данных и методы обработки информации в самоорганизующихся комплексах оперативной диагностики // диссертация на соискание ученой степени доктора технических наук / Москва, 2003
3. Варламов О.О. Миварные технологии как некоторые направления искусственного интеллекта // Проблемы искусственного интеллекта. 2015. № 1 (1). С. 23-37.
4. Варламов О.О. Миварный подход к разработке интеллектуальных систем и проект создания мультипредметной активной миварной интернет-энциклопедии // Известия Кабардино-Балкарского научного центра РАН. 2011. № 1-1 (39). С. 55-64.
5. Варламов О.О. 18 примеров миварных экспертных систем. М.: ИНФРА-М, 2021. 630с.
6. Варламов О.О. Автоматизация умственной деятельности людей через логический искусственный интеллект как фундаментальный механизм развития или гибели человечества // Проблемы искусственного интеллекта. 2017. № 3 (6). С. 23-31.
7. Варламов О.О. Разработка метода распараллеливание потокового множественного доступа к общей базе данных в условиях недопущения взаимного искажения данных // Информационные технологии. 2003. № 1. С. 20-28.
8. Варламов О.О. Миварные базы данных и правил. М.: ИНФРА-М, 2021.351 с.
9. Варламов О.О. Основы создания миварных экспертных систем. М.: ИНФРА-М, 2021. 267 с