Дата подання: 13.09.2022

Дата прийняття:

**УДК 621.391**

**Автоматическое построение нечёткой MISO системы при помощи матричного представления правил и анализ правильности полученной системы.**

**Д. І. Єгошкін, Н. А. Гук**

**AUTOMATIC CONSTRUCTION OF A FUZZY MISO SYSTEM USING A MATRIX REPRESENTATION OF THE RULES AND CORRECTNESS ANALYSIS OF THE RESULTING SYSTEM.**

**Danylo Yehoshkin, Natalia Guk**

*Предлагается алгоритм автоматического формирования нечетких продукционных правил на основании лингвистических переменных и терм-множеств. Целью работы является автоматическое построение базы знаний экспертной системы при помощи объектов обучающей выборки для решения задачи классификации. Для нечёткой базы знаний используется продукционная модель представления знаний, объединяющая модели Мамдани и Такаги-Сугено-Канг. Информация о предметной области представляется в виде продукционной модели в предположении, что левые части правил продукций описывают сочетания признаков объектов, а правые соответствуют классам. Для автоматизации процесса построения правил, антецеденты правил представляются в матричном виде, а консеквенты правил в виде вектора столбца. Для построения матрицы антецедентов предлагается использовать Декартово произведение. Формирование вектора консеквентов осуществляется при помощи процедуры обучения базы знаний. Для анализа правильности базы знаний на корректность используются критерии: полноты, минимальности, связности и непротиворечивости. Для автоматического доказательства полноты базы правил применяется логика Хоара и метод резолюций. Для применения аппарата автоматического доказательства базы правил, будем использовать программное обеспечение Simplify. Simplify – программа, в которой реализован метод резолюций для доказательства истинности заданных предикатов на основании логики первого порядка. Для анализа результата работы классификации экспертной системы используется набор видовой популяции арктических пингвинов. Для оценки качества полученной нечёткой базы правил в работе используются следующие метрики: accuracy, precision, recall, f1-score. В работе проведен анализ результатов классификации и качества полученной базы знаний экспертной системы.*

*Ключевые слова: искусственный интеллект, нечеткая логика, классификация, база правил, экспертная система, база знаний, метод автоматического формирования правил, предикаты первого порядка,* логика Хоара*,* метод резолюций*, Simplify.*

**1. Вступление**

В последнее время подходы к машинному обучению оказали большое влияние на многие отрасли. Развитие технологий позволило широко применять системы искусственного интеллекта для поддержки принятия решений в таких областях знаний как экономика, промышленность, медицина, наука, торговля, строительство, транспорт. Для решения задач управления, идентификации, моделирования сложных физических явлений, классификации, распознавания образов и т.д. [Recent developments and new directions in soft computing. L.A. Zadeh, A.M. Abbasov, R.R. Yager, S.N. Shahbazova, M.Z. Reformat, Eds. STUDFUZ 317, Cham: Springer, 2014. 466 р. DOI 10.1007/978-3-319-06323-2.]. Информация в этих областях знаний представляет собой большие массивы данных, которые необходимо обработать для правильного принятия решений. Обработка потенциально огромного количества данных становится все более актуальной. Для решения широкого класса практических задач сегодня применяются методы искусственного интеллекта: нечеткая логика, нейронные сети, генетические алгоритмы и другие. Это позволяет избежать применения точных математических моделей, основанных на аппарате математических уравнений и классической теории принятия решений. С помощью искусственного интеллекта создаются системы для имитации деятельности экспертов в различных областях. Довольно часто такие системы создаются именно с помощью нечеткой логики, поскольку она на интуитивном уровне соответствует процессу человеческого рассуждения в условиях нечеткости и неполноты условий задачи. После формирования продукционной системы правил необходимо проверить данную систему на корректность, так как система правил может быть неполной, несвязной и противоречивой, поэтому необходимо разработать методы, которые позволят осуществлять процедуру проверки и доказательства правильности полученной системы правил.

2. Анализ литературных данных и постановка проблемы

Задача автоматического формирования и доказательства корректности продукционной базы правил в экспертных системах, является актуальной, в работах используются разные подходы к решению подобных вопросов.

Для проведения анализа достоверности базы знаний нечеткого логического вывода в работе [М. Ю. Терновой] рассматривается представление системы правил в виде метаграфа. При этом нечеткая база знаний представленная в виде модели Мамдани. После процедуры верификации база правил должна удовлетворять следующим условиям: неизбыточности, лингвистической непротиворечивости, отсутствия зацикливания и лингвистической полноты базы нечетких знаний. Предложено проводить статическую верификацию баз нечетких знаний на основе структуры метаграфа, а также его графического представления, путем нахождения зацикливания в структуре графа.

В [КРИВУЛЯ Г.Ф.] рассматривается вопрос автоматизации проверки на корректность базы знаний продукционных правил. Предлагается квантовая модель кодирования символов многозначного алфавита при помощи кубического исчисления. В качестве алгоритма вывода используется алгоритм Мамдани. Разрабатывается компьютерная программа, которая позволяет выполнить формальную проверку базы продукционных правил на корректность при помощи метода резолюций, но теоретико-множественные операции (пересечение, объединение, дополнение, …) над символами алфавита сводятся к соответствующим поразрядным логическим операциям над их кубитами. К недостаткам данного подхода можно отнести следующие: 1) для проверки на корректность есть ряд ограничений: входные лингвистические переменные должны иметь одинаковое количество термов, а термы разных входных лингвистические переменные должны иметь одинаковые диапазоны; 2) продукционные правила должны быть полностью определены, т.е. любому сочетанию значений термов входных лингвистических переменных соответствует определенное значение термов выходной лингвистической переменной.

Также для представления базы знаний применяются ациклические графы, так в работе [ДНФ\_Adnan Darwiche] – для аккумуляции больших наборов знаний в реальном времени, применяется алгоритм построения ориентированного ациклического графа и его последующая трансляция при помощи TLC – (Target Language Compiler). В работе предлагается проводить проверку базы знаний на корректность на этапе трансляции, при помощи как классических методов с применение ДНФ, КНФ, так и с методом NNF (Negation Normal Form – нормальная форма отрицания), что позволяет выводить пользователю контр примеры, если база знаний является противоречивой или не полной.

В случае если необходимо представить данные в табличном виде, для применения матричных и векторных операций обработки, применяется таблица принятия решений (ТПР) [ТАБЛ\_Atsushi]. Большим преимуществом ТПР является наглядность и удобное представление данных для эсперта в процессе анализа и заполнения базы правил. Ещё одним преимуществом является то, что основными операциями продукционного цикла выступают логические (векторные и матричные) операции, позволяющие получить высокую скорость обработки, что, в свою очередь, позволяет весьма эффективно использовать табличную модель представления знаний. Для проверки базы знаний используется дизъюнктивная нормальная форма, удобна для автоматического доказательства теорем. Процесс доказательства основывается на логике высказываний и логике предикатов.

Для генерации базы правил типа Мамдани в работе [Ю.П. Кондратенко] предлагается определение оптимальных консеквентов на основе использования мультиагентных оптимизационных алгоритмов, в частности муравьиных. Данный подход позволяет эффективно генерировать базы правил в следующих случаях: при недостаточном объеме исходной информации; при достаточно большом количестве правил, для которых составление базы нечетких правил на основе знаний экспертов не всегда эффективно, при различном уровне квалификации экспертов.

Указанные подходы существенно повышают эффективность использования интеллектуальных систем, позволяют формировать и проводить верификацию базы продукционных правил, однако накладывают ограничения на количество термов лингвистических переменных, требуют наличие эксперта для анализа работы верификации, накладывают ограничения на количество логических выводов.

В данной работе рассматривается задача классификации. Для решения поставленной задачи применяется подход, основанный на нечёткой логике.

Информация о предметной области представляется в виде продукционной модели в предположении, что левые части правил продукций описывают сочетания признаков объектов, а правые соответствуют классам.

При формировании левых частей продукций использованы признаки объектов из некоторого конечного (квазиконечного, если допустимо пополнение модели) множества, конъюнкция истинных значений которых и определяет условия применимости продукции. В правой части указываются классы из некоторого конечного допустимого множества классов объектов.

Для обеспечения достоверности логического вывода на основе сформулированной базы знаний необходимо, чтобы база знаний обладала свойствами полноты, минимальности (не избыточности), непротиворечивости и связанности.

В данной работе разрабатывается подход, который обеспечивает проверку этих свойств.

**3. Цель и задачи исследования**

Целью исследования является разработка подхода генерация нечетких продукционных правил на основе обучающей выборки для решения задачи классификации.

Для достижения цели были поставлены следующие задачи:

- Создание продукционной модели представления знаний объединяющей модели Мамдани и Такаги-Сугено-Канг

- Построение фазизфикации и дефазификации для данной продукционной модели

- Представление антецедентов правил в виде матрицы, а консеквентов правил в виде вектора столбца для автоматизации процесса построения правил на основе обучающей выборки.

- Анализ правильности базы знаний на корректность при помощи критериев: полноты, минимальности, связности и непротиворечивости. Для доказательства полноты базы знаний применить логику Хоара, метод резолюций и программу для автоматического доказательства Simplify.

- Применение подхода для решения задачи классификации и оценка качества полученной нечёткой базы правил при помощи следующих метрик: accuracy, precision, recall, f1-score.