

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра математической кибернетики и компьютерных наук

НЕЧЕТКИЙ ЛОГИЧЕСКИЙ ВЫВОД В ЭКСПЕРТНЫХ СИСТЕМАХ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 5 курса 551 группы
направления 09.03.04 — Программная инженерия
факультета КНиИТ
Голикова Леонида Олеговича

Научный руководитель

к. ф.-м. н., доцент

А. С. Иванов

Заведующий кафедрой

к. ф.-м. н., доцент

С. В. Миронов

Саратов 2022

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ	5
1.1 Экспертные системы	5
1.2 Нечеткая логика	7
1.3 Описание реализации экспертной системы	8
1.4 Тестирование базы знаний	9
1.5 Интерфейс	10
ЗАКЛЮЧЕНИЕ	11

ВВЕДЕНИЕ

В настоящее время существует огромный интерес к экспертным системам во многих областях деятельности человека. Экспертные системы используются для решения различных задач в таких областях, как:

- газовая промышленность;
- финансы;
- космос;
- фармакология;
- образование;
- телекоммуникации;
- медицина и др.

Среди всех направлений, использующих экспертные системы, одним из наиболее важных и актуальных является медицина.

В медицине нередко встречается неточность: диагнозов, признаков заболеваний или экспертных оценок врача. Также в медицине существуют следующие проблемы:

- полиморфизм клинических проявлений болезни;
- симптомы-миражи (появление которых ошибочно связывают с патологией определенной системы);
- болезни-хамелеоны (которые маскируются под другие заболевания) и т. д.

В таких условиях, где часто встречается отсутствие точных или закономерных данных, а также неточность экспертных оценок, установка диагноза на основании признаков заболевания в медицинских системах часто бывает затруднительным. Поэтому при разработке экспертных систем для использования в медицине актуально использование технологий, основанных на нечетких вычислениях.

В нечеткой логике, в отличие от классической, вместо величин истина и ложь используется величина степени истинности, принимающая любые значения из бесконечного множества от 0 до 1 включительно. Такой подход позволяет построить более гибкую систему в условиях неточных данных, где точность результата будет зависеть от набора правил нечеткой продукционной базы знаний и определений нечетких высказываний.

Целью работы является реализация учебного проекта, представляющего собой прототип экспертной системы с нечетким логическим выводом для определения диагноза, по указанным пользователем симптомам, а также исследование применимости нечетких экспертных систем в области медицины.

Для достижения цели должны быть решены следующие задачи:

- реализация нечеткой продукционной базы знаний;
- реализация механизма тестирования правил базы знаний;
- реализация алгоритма определения диагноза по симптомам;
- реализация механизма загрузки данных;
- реализация панели для управления данными;
- реализация механизма авторизации пользователей;
- реализация интерфейса.

Методологические основы нечеткой логики и принципов построения экспертных систем представлены в работах: «Нечеткая логика» (авторства Ф. Ленге), «Разработка интеллектуальных систем» (авторства Д. Хултена), «Инженерия знаний. Модели и методы. Учебник для вузов» (авторства Д. Кудрявцева, Т. Гавриловой и Д. Муромцева), «Интеллектуальные системы: нечеткие системы и сети 2-е изд., испр. и доп. Учебное пособие для вузов» (авторства В. Горбаченко и Б. Ахметова) и «Интеллектуальные системы» (авторства В. Иванова).

Практическая значимость бакалаврской работы заключается в возможности использования реализованного прототипа экспертной системы в качестве основы для полноценной версии приложения, решающего более широкий спектр задач. Значимость работы также обусловлена широкой востребованностью исследуемых практик и высокой актуальностью подобных систем в последнее время.

Бакалаврская работа состоит из введения, 2 разделов, заключения, списка использованных источников и одного приложения. Общий объем работы – 76 страниц, из них 40 страниц – основное содержание, включая 20 рисунков, цифровой носитель в качестве приложения, список использованных источников информации – 20 наименований.

1 КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Теоретическая часть» посвящен теоретическим основам и истории развития нечетких вычислений и экспертных систем с нечетким логическим выводом.

1.1 Экспертные системы

Экспертная система — это программа (или набор программ), которая выполняет функции эксперта при решении какой-либо задачи в области его компетенции. Такие системы представляют собой отдельный класс интеллектуальных систем, т. е. систем, способных решать задачи, традиционно считающиеся творческими и принадлежащие конкретной предметной области. Экспертные системы предназначены для так называемых неформализованных (неточных) задач, т. е. экспертные системы не отрицают и не заменяют традиционного подхода к разработке программ, ориентированного на решение формализованных (точных) задач.

Разработка экспертных систем имеет существенные отличия от разработки обычного программного продукта. При разработке экспертных систем, как правило, используется концепция «быстрого прототипа». Суть этой концепции состоит в том, что разработчики не пытаются сразу построить конечный продукт. На начальном этапе они создают прототип системы, который должен удовлетворять двум требованиям: быть способным решать типовые задачи и иметь незначительные затраты времени на разработку, чтобы можно было реализовать параллельно процесс накопления и отладки знаний, осуществляемый экспертом и процессом разработки системы инженером по знаниям и программистом. Для удовлетворения этим требованиям при создании прототипа, как правило, используются разнообразные технологии, ускоряющие процесс проектирования. В случае успеха эксперт с помощью инженера по знаниям расширяет знания прототипа о проблемной области.

В ходе работ по созданию экспертных систем сложилась определенная технология их разработки, включающая шесть следующих этапов: идентификация, концептуализация, формализация, реализация, тестирование, опытная эксплуатация и внедрение.

На этапе идентификации определяются задачи, которые будет решать система, выявляются цели разработки, определяются эксперты и типы пользова-

телей. Определение участников и их ролей сводится к определению количества экспертов и инженеров по знаниям, а также формы их взаимоотношений.

На этапе концептуализации проводится содержательный анализ проблемной области, определяются используемые понятия, их взаимосвязи и методы решения задач.

Этап формализации определяет способы представления всех видов знаний, на нем формализуются основные понятия, моделируется работа системы, оценивается соответствие зафиксированных понятий целям системы, производится определение способов представления и манипулирования базой знаний. Основными задачами в процессе формализации являются структуризация исходной задачи и необходимых знаний.

На этапе реализации осуществляется наполнение экспертом базы знаний. В связи с тем, что основой экспертной системы являются знания, данный этап является наиболее важным и трудоемким в разработке системы.

Этап тестирования экспертной системы рассматривают в качестве заключительной фазы процесса разработки. Выделяют три аспекта тестирования экспертных систем: тестирование исходных данных, логическое тестирование базы знаний, концептуальное тестирование прикладной системы.

Тестирование исходных данных включает проверку информации, служащей основой для проведения экспертизы. Очевидно, что наборы данных, используемых при тестировании, должны покрывать область возможных ситуаций, распознаваемых экспертной системой.

Логическое тестирование базы знаний заключается в обнаружении логических ошибок, не зависящих от предметной области, таких, как: избыточные, циклические и конфликтные правила, пропущенные и пересекающиеся правила, несогласуемые и терминальные условия. Характер этих ошибок позволяет автоматизировать процесс логического тестирования.

Концептуальное тестирование проводится для проверки общей структуры системы и учета в ней всех аспектов решаемой задачи. На этом этапе проведение тестирования невозможно без привлечения конечных пользователей системы.

На этапе опытной эксплуатации и внедрения проверяется пригодность экспертной системы для конечного пользователя.

Интеллектуальные системы, основанные на использовании математического аппарата нечеткой логики, являются наиболее простыми представителя-

ми систем искусственного интеллекта. Их также принято называть термином «нечеткие системы». Отличительной особенностью нечетких систем, по сравнению с другими классами интеллектуальных систем, является использование правил логического вывода, написанных человеком. Данные правила относятся к простейшим правилам нечеткой продукционной базы знаний, используемым в экспертных системах и имеют вид «Если-то».

Нечеткие системы используются тогда, когда:

- неопределенность наших знаний о рассматриваемом объекте носит сравнительно простой и преимущественно количественный характер и может быть описана с помощью функции принадлежности (аналогов функции распределения в теории вероятности);
- возможно применение для формирования управляющих воздействий (команд) математически строго определенных правил нечеткого вывода (аналогично известным правилам логического вывода в булевой алгебре);
- используется достаточно простая словесная или лингвистическая форма представления управляющих воздействий ограниченного потока слов и предложений словесного языка, вида «Если-то» (условие - вывод).

1.2 Нечеткая логика

Нечеткая логика — это форма многозначной логики, в которой истинные значения переменных могут быть любыми действительными числами от 0 до 1 включительно в то время, как в классической (булевой) логике истинные значения переменных могут быть только значениями 0 или 1.

Нечеткая логика используется в концепции «частичной истины». Это дает возможность определять понятия, нечеткие по своей природе: «хороший», «высокий», «слабый» и т. д.

Процесс нечеткого логического вывода представляет собой некоторый алгоритм получения нечетких заключений на основе нечетких условий. Основой для проведения операции нечеткого логического вывода является база правил, содержащая нечеткие высказывания в форме «Если-то» и функции принадлежности для соответствующих лингвистических термов.

На сегодняшний день особенно распространены два алгоритма нечеткого логического вывода: алгоритм Сугено и алгоритм Мамдани. Их основное отличие заключается в том, что алгоритм Сугено выдает четкий (количественный) результат в виде значения линейной функции, а алгоритм Мамдани —

качественный результат (нечеткую переменную).

Таким образом в теоретической части бакалаврской работы рассмотрены основные понятия, необходимые для понимания экспертных систем с нечетким логическим выводом, а также принципы разработки таких систем.

Второй раздел «Практическая часть» посвящен реализации прототипа экспертной системы с нечетким логическим выводом для определения диагноза по указанным симптомам.

1.3 Описание реализации экспертной системы

В качестве нечетких переменных в программе использованы симптомы в условной части правил и названия болезней в части заключений.

Алгоритм Сугено был выбран в качестве основы для реализованного алгоритма определения диагноза на основе симптомов.

В приложении выбор функции принадлежности реализован автоматически и зависит от положения значения симптома в диапазоне значений симптома.

При разработке экспертной системы была использована справочная информация из медицинской литературы, а также ресурса, расположенного по адресу msdmanuals.com, который содержит справочные материалы для медицинских работников с подробными описаниями болезней.

Основной сценарий использования приложения выглядит следующим образом. Пользователь выбирает набор симптомов и процентное значение их выраженности на форме ввода входных данных. После отправки формы запустится процесс определения диагноза на основе базы знаний системы, результатом которого будут являться 3 диагноза с наивысшей степенью истинности. На форме вывода по каждому результату будут отображены: название диагноза, его степень истинности, а также справочная информация.

Был разработан набор из 30 правил нечеткой продукции, который охватывает 20 различных диагнозов, большинство из которых классифицируются, как вирусы и бактериальные инфекции. Каждое правило состоит из:

- набора условий (conditions), состоящих из:
 - набора подусловий (subconditionsOperationType), которые состоят из:
 - * симптома (symptom);
 - * терма (term).
 - логической операции над подусловиями (subconditionsOperationType).
- операции над условиями правила (conditionsOperationType);

- весового коэффициента (weight);
- названия болезни (disease).

В правилах системы предусмотрены 2 значения типов логических операции: конъюнкция (AND) и дизъюнкция (OR).

1.4 Тестирование базы знаний

В экспертной системе было реализовано автоматическое логическое тестирование базы правил. Для этих целей был реализован консольный модуль тестирования.

Сначала производится проверка на дубликаты правил в базе данных.

После успешной проверки на дубликаты по каждому диагнозу из выборки запускается основной алгоритм приложения, с использованием всех симптомов из подусловий каждого правила, в котором присутствует диагноз. Значения на вход алгоритма Сугено выбираются из диапазона значений термов в составе подусловий выборки.

На каждой итерации производятся следующие проверки результата:

- диагноз присутствует в результате;
- степень истинности диагноза больше 50%;
- степень истинности диагноза самая большая среди других диагнозов в результате.

После завершения последней итерации тестирования, результаты по каждому диагнозу выводятся в консоль.

В качестве базы знаний использована база данных MySQL. Базовый набор данных для демонстрации работы системы хранится в виде JSON файла конфигурации. Для загрузки, обработки и перемещения данных из файла конфигурации в базу данных реализована отдельная консольная команда. При выполнении команды загрузки данных таблицы базы данных очищаются, а данные из файла конфигурации проверяются на ошибки и записываются в базу данных. В команде парсинга реализована подробная нотификация об ошибках, которая выводится в консоль и позволяет быстро отредактировать файл конфигурации

В приложении реализована авторизация пользователей. Существуют две роли пользователей в системе: гость и авторизованный пользователь. Гость не имеет прав редактирования базы данных экспертной системы, в отличие от авторизованного пользователя.

1.5 Интерфейс

Вспомогательные функций экспертной системы были реализованы в виде команд с консольным интерфейсом, а основной функционал системы был реализован в веб-интерфейсе.

Интерфейс приложения выполнен с использованием стандартного на сегодняшний день набора веб-технологий:

- язык гипертекстовой разметки HTML;
- язык программирования JavaScript;
- язык программирования PHP;
- каскадные таблицы стилей CSS;
- библиотека Bootstrap;
- шаблонизатор Laravel Blade.

Дизайн большинства страниц экспертной системы был реализован с использованием библиотеки Bootstrap и шаблонизатора Laravel Blade в качестве средств для ускорения реализации системы, в соответствии с рекомендациями по разработке экспертных систем.

В практической части работы подробно изображена каждая страница приложения, а также примеры вывода разработанных консольных команд.

Итоговая экспертная система выполняет свою основную задачу, имеет удобный веб-интерфейс, а также имеет механизмы манипуляции данными и тестирования базы правил с подробными уведомлениями об ошибках.

ЗАКЛЮЧЕНИЕ

В результате бакалаврской работы был реализован учебный проект, являющийся прототипом экспертной системы с нечетким логическим выводом для определения диагноза по указанным пользователем симптомам.

В прототипе экспертной системы были реализованы:

- алгоритм определения диагноза по симптомам;
- механизм загрузки данных с командным интерфейсом;
- механизм тестирования правил базы знаний;
- панель управления данными;
- механизм авторизации пользователей;
- нечеткая продукционная база знаний;
- веб-интерфейс приложения.

Реализованная экспертная система продемонстрировала применимость принципов нечетких экспертных систем для областей деятельности с нечеткими данными на примере медицины. Точность логических выводов таких систем напрямую зависит от качества и объема нечеткой продукционной базы знаний. Основные источники информации:

1. Ланге, Ф. Нечеткая логика / Ф. Ланге. — СПб.: Страта, 2018. — С. 130.
2. Хултен, Д. Разработка интеллектуальных систем / Д. Хултен. — М.: ДМК Пресс, 2020. — С. 284.
3. Кудрявцев, Д. Инженерия знаний. Модели и методы. Учебник для вузов / Д. Кудрявцев, Т. Гаврилова, Д. Муромцев. — СПб.: Лань, 2022. — С. 325
4. Горбаченко, В. Интеллектуальные системы: нечеткие системы и сети 2-е изд., испр. и доп. Учебное пособие для вузов / В. Горбаченко, Б. Ахметов. — М.: Юрайт, 2018. — С. 105.
5. Иванов, В. Интеллектуальные системы / В. Иванов. — М.: Юрайт, 2022. — С. 94.