

ОТЗЫВ

официального оппонента

на диссертационную работу Натальи Владимировны Сапкиной
«Восстановление закономерностей на основе нечетких
регрессионных моделей», представленную на соискание ученой
степени кандидата технических наук по специальности
05.13.17 – «Теоретические основы информатики»

1. Актуальность темы диссертационной работы

В настоящее время происходит бурное развитие методов гибридных интеллектуальных систем, что невозможно без решения задач, связанных с накоплением, обработкой и хранением информацией для извлечения знаний. Особенно это актуально для интеллектуального анализа данных, позволяющего выявить полезную скрытую информацию для принятия решения в форме зависимостей определенного вида. Заметим, что исходная информация, используемая для восстановления закономерностей, может быть приближенной, что делает невозможным применение классических методов для ее обработки. Целесообразным, на наш взгляд, является адаптация существующих моделей и методов. Один из наиболее широко применяемых подходов к восстановлению зависимостей в выборочных данных основан на методе регрессионного моделирования, который относится к статистическим методам, при этом обязательным условием его применения является принадлежность данных к множеству действительных чисел, что существенно снижает область его практического применения для интеллектуального анализа данных. Развитие методов нечеткого регрессионного моделирования, ориентированного на нечеткие числа, произошло сравнительно недавно и любое исследование, расширяющее положения данного подхода, уже само по себе актуально. Построение нечетких моделей опирается на математический аппарат, включающий определение арифметических операций над нечеткими числами и их сравнение. Однако не для всех типов нечетких чисел результат арифметической операции представляет собой нечеткое число того же типа,

что требует дополнительных исследований. Возникает актуальная задача совершенствования существующих методов нечеткого регрессионного моделирования для восстановления закономерностей в данных путем учета различных типов данных и параметров модели, представленных нечеткими числами L - R -типа.

Решению этой задачи, предметно ориентированной на разработку инструментальных средств нечеткого регрессионного моделирования в виде математического, алгоритмического, и программного обеспечения информационной системы анализа и интеллектуальной обработки данных, посвящена диссертационная работа Н.В. Сапкиной, что определяет актуальность темы исследования и значимость для теории и практики полученных результатов.

2. Достоверность и новизна основных выводов и результатов диссертации

Анализ содержания диссертации позволяет сделать вывод, что задачи исследования в основном решены. Основные результаты диссертационной работы являются новыми научными разработками. Их новизна заключается в следующем.

1. Исследованы алгебраические свойства арифметических операций над нечеткими числами L - R -типа, введено понятие группоида с одной или двумя операциями, позволяющего в зависимости от выполняемых свойств определить тип алгебры с одной или двумя операциями. Показано, что группоид с операцией сложения не имеет обратных элементов. Найден обратный элемент и определены условия его существования для группоида с операцией умножения. Совокупность перечисленных теоретических результатов, на мой взгляд, является наиболее весомым научным результатом, поскольку позволяет осуществлять вычисления при построении нечетких регрессионных моделей.

2. Разработан метод оценки параметров нечетких линейных (парной и множественной) регрессионных моделей для задач, в которых исходная информация является полностью или частично приближенной, а коэффициенты моделей представлены обычными и/или нечеткими числами, что позволяет модифицировать общую процедуру регрессионного моделирования для восстановления закономерностей в разнородных и приближенных данных на основе использования нечетких чисел *L-R*-типа.

3. Предложены подходы к выявлению множества существенных переменных, которые делятся на две группы. Первая группа обобщает классические статистические методы (коэффициент корреляции, стандартизированное уравнение множественной линейной регрессии) на случай, когда исходная информация является приближенной и представлена нечеткими числами. С помощью операций над нечеткими числами выведен нечеткий коэффициент корреляции и коэффициенты стандартизированного уравнения регрессии. Ко второй группе относится метод выявления существенных переменных, основанный на применении автоассоциативных нейронных сетей. Поскольку обучающая выборка представлена в виде кортежей нечетких чисел, то процедура обучения Ойя была адаптирована для работы с приближенной информацией за счет полученных формул для выходных значений и весовых коэффициентов.

4. Разработана структурная модель информационной системы интеллектуального анализа данных, основанная на технологии информационного хранилища, включающая в качестве инвариантной составляющей блок нечеткой арифметики, который может использоваться как самостоятельное приложение, а также проблемно-ориентированную составляющую, содержащую модули для проведения нечеткого линейного регрессионного моделирования.

Следует заметить, что основой проведенного исследования является общая процедура регрессионного моделирования, и в диссертации рассмотрены практически все ее этапы, в том числе, предложены подходы к оценке точности и адекватности нечетких регрессионных моделей.

В целом основные результаты диссертации представляются научно-обоснованными и достоверными. Этот факт обусловлен, с одной стороны, корректным использованием методологической основы исследования, которую составляют методы математической статистики, регрессионного и нейросетевого моделирования, линейная алгебра, теория нечетких множеств и нечеткая арифметика, а с другой – взаимно дополняющие друг друга положения и результаты теоретического исследования и проведенного вычислительного эксперимента. Достоверность научных положений диссертационной работы, правильность конкретных выводов и рекомендаций, подтверждены созданием работоспособных инструментальных средств в виде алгоритмического и программного обеспечения предметной информационной системы анализа и интеллектуальной обработки данных, с помощью которой получены положительные результаты экспериментов по восстановлению закономерностей в данных на основе нечеткого линейного регрессионного моделирования. Полученные выводы экспериментального исследования разработанных нечетких регрессионных моделей соответствуют существующим классическим результатам статистического моделирования закономерностей, а также принципу декомпозиции теории нечеткого моделирования, согласно которому всякая модель с нечеткими коэффициентами может быть представлена в виде совокупности обычных (четких) моделей с коэффициентами в виде α -срезов соответствующих нечетких коэффициентов.

Апробация исследований в научной печати и на конференциях различного ранга также подтверждает корректность выводов и правильность полученных в диссертации результатов.

Результаты диссертационной работы Н.В. Сапкиной внедрены в учебный процесс Воронежского государственного университета и в практическую деятельность ЗАО ЛЦ «АВС Фарбен» (г. Воронеж), что подтверждено соответствующими актами.

3. Ценность диссертационной работы для науки и практики

Значимость результатов диссертационной работы Сапкиной Н.В. для науки заключается в том, что разработанные ею методы восстановления закономерностей в данных теоретически обобщают и развивают подходы классического линейного регрессионного моделирования на случай приближенной информации при условии, что коэффициенты моделей могут быть как четкими, так и нечеткими. Теоретические результаты, касающиеся существования алгебраических структур на множестве нечетких чисел позволят корректно выполнять алгебраические преобразования при построении нечетких моделей в различных прикладных областях.

Работа вносит также вклад в развитие методов обработки информации (больших объемов) с целью выявления существующих закономерностей и формирования знаний. В качестве модели знаний может выступать, например, продукционная модель, когда в посылке продукционного правила информация задается приближенно с помощью лингвистических переменных, а заключение содержит функцию регрессии, которая описывает зависимость выходной переменной от входных.

Практическое значение диссертационной работы Н.В. Сапкиной заключается в создании комплекса инструментальных средств в виде предметно-ориентированных моделей, алгоритмов и пакета прикладных программ, которые, по сути, образуют прототип информационной системы интеллектуального анализа данных, основанной на технологии информационного хранилища и предназначенной для автоматизированного поиска закономерностей и скрытых правил в базе данных, что позволяет на

основе процедур анализа и прогнозирования оптимизировать процессы принятия решений при управлении социально-экономическими и другими системами.

4. Замечания по диссертации

1) Глава 1 содержит только описание существующих теоретических результатов без указания тех проблем, которые являются предметом исследования. В диссертации не отражены сравнительные характеристики полученных результатов с результатами аналогичных исследований других авторов.

2) Во введении при описании объекта и предмета исследования указывается свойство линейности регрессионных моделей, однако в формулировке цели работы данное свойство моделей отсутствует.

3) В диссертации предложено несколько подходов для выявления существенных переменных, но отсутствуют рекомендации по выбору.

4) Алгоритм, представленный на рис.3.2, оформлен не по ГОСТу (не показаны начало и останов вычислений), а схема на рис. 4.5 – зациклена.

Перечисленные замечания по диссертации не влияют на общее впечатление о работе. Диссертация написана грамотно, текст изложен четким и ясным языком, оформление соответствует требованиям ВАК. Автореферат соответствует содержанию диссертационной работы. По результатам проведенных исследований Н.В. Сапкиной опубликовано 14 научных работ, из них 5 – из перечня изданий, рекомендованных ВАК РФ, 7 опубликованы в материалах конференций, имеется 2 свидетельства о государственной регистрации программ, что позволяет оценить уровень публикаций как более, чем достаточный.

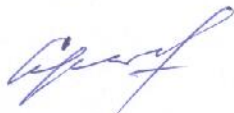
5. Заключение

Диссертационная работа Сапкиной Натальи Владимировны «Восстановление закономерностей на основе нечетких регрессионных

моделей» является законченной научно-квалификационной работой, содержит результаты, характеризующиеся научной новизной, в области нечеткого регрессионного моделирования, удовлетворяет требованиям «Положения о присуждении ученых степеней», предъявляемым к кандидатским диссертациям.

Считаю, что автор диссертации, Сапкина Наталья Владимировна, заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики».

Официальный оппонент,
профессор кафедры вычислительной техники
и информационных систем Воронежской
государственной лесотехнической академии,
докт. техн. наук, профессор



Сербулов Юрий Стефанович

userbulov@vglta.vrn.ru
т. 8-910-240-71-33,
394613, г.Воронеж, ул.Тимирязева, д.8, ВГЛТА



Я подтверждаю подпись *Сербулов Ю. С.*
удостоверяю: *Сербулов Ю. С.*
подпись ректората *Сербулов Ю. С.*
20.05.20 г.