

ОТЗЫВ

официального оппонента

Фридмана Александра Яковлевича

на диссертационную работу Шкабериной Гузели Шарипжановны
«Модели и алгоритмы автоматической классификации продукции»,
представленной на соискание ученой степени кандидата технических наук по
специальности 05.13.01 – «Системный анализ, управление и обработка
информации (космические и информационные технологии)»

Актуальность темы диссертационной работы

Работа посвящена исследованию и разработке алгоритмов автоматической классификации продукции, которые позволяют повысить точность и стабильность результата решения практических задач. Методы автоматической группировки являются частью машинного обучения, актуальность которого возрастает с каждым годом. Поиск алгоритма автоматической классификации объектов, обладающего одновременно высокой точностью, стабильностью результата, и при этом высокой скоростью работы, является одной из проблем автоматической классификации объектов. Проведенный Шкабериной Г.Ш. анализ существующих подходов решения задач автоматической группировки продукции выявил дефицит таких алгоритмов.

Представленные в работе модель и алгоритм автоматической группировки продукции на основе модели k -средних с расстоянием Махаланобиса и средневзвешенной ковариационной матрицей, генетический алгоритм с перекрестной мутацией для задачи k -средних позволяют повысить точность автоматической классификации промышленной продукции и получить более стабильные значения целевой функции. А предложенный алгоритм обучения двухслойной сигмоидальной искусственной нейронной сети с регуляризацией демонстрирует более высокую точность классификации продукции. Предложенные алгоритмы и модели вносят существенный вклад в решение задач автоматической классификации объектов.

В связи с изложенным, тему диссертационной работы Шкабериной Гузели Шарипжановны следует признать актуальной.

Структура диссертации

Диссертационная работа Шкабериной Г.Ш. с приложениями представлена на 222 страницах, основной текст состоит из введения и четырех глав.

Во введении обосновывается актуальность диссертационного исследования, формулируется цель и основные задачи, характеризуется степень новизны полученных результатов и их апробация.

В первой главе проведен обзор существующих методов решения задач автоматической группировкой объектов, выделены проблемы существующих методов. Интересной представляется проведенная аналогия между задачами классификации и размещения.

Во второй главе исследовано практическое применение факторного анализа для решения задачи автоматической группировки продукции, исследована возможность применения жадных эвристических алгоритмов для задач разделения с различными мерами расстояний, также представлена модель и алгоритм для решения задачи автоматической группировки продукции на основе модели k -средних с мерой расстояния Махаланобиса и средневзвешенной ковариационной матрицей.

Третья глава посвящена разработке генетического алгоритма с перекрестной мутацией для задачи k -средних. Новый подход показал, что генетические алгоритмы с перекрестной мутацией превосходят генетические алгоритмы без какой-либо процедуры мутации и алгоритмы с равномерной случайной мутацией.

В четвертой главе представлен алгоритм обучения двухслойной сигмоидальной нейронной сети с регуляризацией с новым подходом к нахождению начального приближения.

Достоверность и обоснованность научных положений, выводов и рекомендаций

Автором работы проведен подробный анализ отечественной и зарубежной литературы по кластерному анализу и генетическим алгоритмам. Представленные в работе положения и выводы достаточно обоснованы численными экспериментами. Достоверность результатов работы подтверждается их публикацией в рецензируемых научных изданиях. Результаты прошли всестороннюю апробацию на конференциях и используются в деятельности АО «Испытательный технический центр – НПО ПМ».

Научная новизна результатов, ценность для науки

Научная ценность работы состоит в том, что автор предлагает новую модель и алгоритм автоматической группировки объектов, основанный на оптимизационной модели k -средних с мерой расстояния Махаланобиса со средневзвешенной ковариационной матрицей, что позволяет снизить долю ошибок. Также разработан новый генетический алгоритм с перекрестной мутацией для задачи k -средних, который позволяет, по сравнению с известными алгоритмами автоматической группировки, статистически значимо повысить точность результата, а также его стабильность за фиксированное время. Предложен алгоритм обучения двухслойной сигмоидальной искусственной нейронной сети с регуляризацией, демонстрирующий более высокую точность классификации продукции.

Представленные в работе результаты получены лично автором и являются новыми.

Научная значимость работы состоит в том, что предложенные алгоритмы дополняют список эффективных методов решения задач кластеризации.

Ценность результатов работы заключается в новом подходе к решению задач кластеризации и классификации с повышенными требованиями к точности и стабильности результата, обусловленными широким диапазоном

сфер их применения, например, при проверке качества промышленных изделий на производстве.

Общая оценка работы

Диссертация изложена на 222 страницах, состоит из введения, четырех глав, заключения, списка литературы и приложений. Содержание работы и представленные результаты соответствуют п. 4 паспорта специальности 05.13.01 – Системный анализ, управление и обработка информации.

Выявленные недостатки и замечания

1. Не проанализированы возможности решения задачи диссертации с помощью аппарата мультимножеств (работы А.Б. Петровского).
2. Перевод термина «fitness function» на стр. 25 как «функция полезности» вызывает сомнения.
3. На стр. 41 есть нерасшифрованная аббревиатура ЭРИ.
4. Во второй главе результаты кластеризации микросхем 1526ТЛ1 и 1526ИЕ10 и диодов 3ОТ122А, представленные в таблицах (стр. 91-92), для наглядности, можно было отобразить в виде рисунков (например, гистограмм).
5. В третьей главе на стр. 169 описан набор данных SUSY. Указано, что не учитывается истинная маркировка данных. В таблице 3.7 представлены результаты вычислительного эксперимента над этими данными. Было бы интересно сравнить истинную маркировку с полученным результатом.
6. В третьей главе часть алгоритмов можно было представить в виде блок-схем для наглядности.
7. На стр. 126 перед таблицей 3.2 в конце последнего абзаца по тексту указан уровень значимости 0,99, то же самое указано и на стр. 130. Скорей всего, предполагается $p < 0,01$.
8. Проведен обширный анализ аналогов и прототипов выполненных исследований, выполнены многочисленные расчеты и эксперименты по различным этапам решения поставленной задачи, но не приведена блок-схема или описание общей последовательности предложенного решения задачи работы для различных возможных составов партий продукции.
9. Одни и те же обозначения (в частности, количество групп k) пишутся то курсивом, то обычным шрифтом).
10. В тексте диссертации и автореферата имеются многочисленные грамматические и синтаксические ошибки.

Заключение о соответствии диссертации требованиям и критериям, установленным Положением о порядке присуждения ученых степеней

Вышеуказанные замечания не снижают общей положительной оценки работы. Диссертация представляет собой завершенное научное исследование на актуальную тему, выполненное самостоятельно и на достаточно высоком научном уровне. Полученные результаты достоверны, имеют теоретическую и практическую значимость. Область научных исследований и тематика

диссертационной работы соответствуют заявленной специальности, выводы и заключения, сделанные автором, обоснованны.

Автореферат соответствует основному содержанию диссертации, содержит описание основных этапов исследования, полученные результаты и выводы. Оформление автореферата и диссертации соответствует требованиям ВАК РФ.

Представленная диссертационная работа отвечает требованиям п. 9 Положения о присуждении ученых степеней и постановления Правительства РФ от 24 сентября 2013, № 842 «О порядке присуждения ученых степеней», предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор, Шкаберина Гузель Шарипжановна, заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (космические и информационные технологии)

Официальный оппонент,
доктор технических наук, профессор,
ведущий научный сотрудник
лаборатории информационных
технологий управления
промышленно-природными
системами Федерального
государственного бюджетного
учреждения науки Институт
информатики и математического
моделирования Федерального
Исследовательского Центра
«Кольский Научный Центр
Российской академии наук» (ИИММ
КНЦ РАН)

Фридман Александр Яковлевич

Адрес организации:
184209 Мурманская область,
город Апатиты, улица Ферсмана, дом 24А
e-mail: alex.ya.fridman@gmail.com

14.01.2021

Подпись доктора технических наук, профессора Фридмана Александра Яковлевича заверяю
Ученый секретарь ИИММ КНЦ РАН,
кандидат технических наук

И.О. Датьев

