

УТВЕРЖДАЮ

Врио директора Федерального государственного бюджетного научного учреждения «Федеральный исследовательский центр «Красноярский научный центр Сибирского отделения Российской академии наук»
доктор сельскохозяйственных наук



А.А. Шпедт

2021 г.

ОТЗЫВ

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

Актуальность исследований. Нарастающие объемы цифровых данных во всех сферах современного общества обуславливают необходимость развития новых эффективных методов интеллектуального анализа. Диссертационное исследование Г.Ш. Шкабериной посвящено развитию методов автоматической группировки объектов с целью повышения эффективности и точности решения практических задач в различных областях, в том числе для выделения групп однородной продукции в космическом приборостроении, при массовом производстве радиоизделий и в других отраслях, где требуется классификация продукции с повышенными требованиями к качеству. Тема диссертационного исследования является актуальной.

Цель и задачи исследования. Целью диссертационной работы является повышение точности и стабильности результатов решения задач автоматической группировки объектов. Для достижения поставленной цели автором решены следующие основные задачи.

1. Построение модели автоматической группировки объектов на основе модели k -средних с расстоянием Махаланобиса с применением методов

факторного анализа для предварительного снижения размерности исходных данных, позволяющих снизить долю ошибок автоматической группировки промышленной продукции по сравнению с известными моделями.

2. Разработка генетического алгоритма для задачи k -средних с перекрестной мутацией, позволяющего повысить точность и стабильность решения по достигаемому значению целевой функции за фиксированное время выполнения по сравнению с известными алгоритмами автоматической группировки объектов.

3. Разработка алгоритма обучения двухслойной сигмоидальной искусственной нейронной сети с регуляризацией, повышающего точность решения задачи классификации промышленных изделий на однородные производственные партии при наличии малоинформативных признаков.

Научная значимость работы заключается в разработке новых и модификации существующих моделей, методов и алгоритмов, обеспечивающих повышение точности и стабильности результатов решения задач автоматической группировки объектов. Принцип использования единой процедуры в качестве оператора скрещивания и мутации создает основу для синтеза новых эффективных алгоритмов решения широкого круга NP-трудных задач.

Новые научные результаты диссертационной работы:

1. Предложена новая модель для решения задач автоматической группировки промышленной продукции на основе модели k -средних с расстоянием Махаланобиса с применением метода главных компонент. Применение новой модели позволяет повысить точность решения (индекс Рэнда) для задачи выделения однородных производственных партий изделий по данным тестовых испытаний.

2. Предложен новый алгоритм автоматической группировки объектов, основанный на оптимизационной модели k -средних с мерой расстояния Махаланобиса и средневзвешенной ковариационной матрицей, рассчитанной по обучающей выборке. Алгоритм позволяет снизить долю ошибок (повысить индекс Рэнда) при выявлении однородных производственных партий продукции по результатам тестовых испытаний.

3. Разработан новый генетический алгоритм для задачи k -средних с применением единой жадной агломеративной эвристической процедуры в качестве оператора скрещивания и оператора мутации. По сравнению с известными алгоритмами автоматической группировки применение данного алгоритма позволяет статистически значимо повысить стабильность и точность результата, улучшив достигаемое значение целевой функции в рамках выбранной математической модели решения задачи автоматической группировки за фиксированное время.

4. Разработан новый алгоритм обучения двухслойной сигмоидальной искусственной нейронной сети с регуляризацией, демонстрирующий более высокую точность классификации промышленной продукции по данным тестовых испытаний в сравнении с методами обучения таких нейронных сетей при известных методах регуляризации.

Практическая значимость работы. Результаты диссертационного исследования имеют высокую практическую значимость, так как могут применяться для задач разделения объектов как при производстве и тестировании продукции, так и в других отраслях, где требуется классификация изделий с повышенными требованиями к качеству. Использование предложенных алгоритмов решения задач автоматической группировки и классификации в различных системах, предназначенных для решения таких задач за фиксированное время, позволяет существенно повысить точность получаемых решений в рамках выбранной модели по получаемому значению целевой функции.

Автором представлено свидетельство о государственной регистрации программы для ЭВМ и акт о внедрении результатов диссертационной работы. Рекомендуются использование результатов диссертационной работы для решения задач автоматической группировки продукции в АО «Испытательный технический центр-НПО ПМ», в АО «Информационные спутниковые системы» им. академика М.Ф. Решетнева, в АО «Научно-производственное предприятие «Радиосвязь» и на других предприятиях, осуществляющих производство и применение больших объемов промышленной продукции с повышенными требованиями к ее качеству.

Замечания по содержанию диссертации.

1. К сожалению, анализ известных алгоритмов автоматической классификации проведён не полно, отсутствует их типизация. В работах А.А. Дорофеюка выделяются три группы автоматической классификации: основанные на знании, либо принятии гипотез о мере близости между объектами классификации; требующие знания цели классификации в виде некоторого функционала; предполагающие явное определение класса. Предлагаемые автором методы классификации относятся к первой группе алгоритмов и частично ко второй. Признание этого факта позволило бы более точно сформулировать название работы, её цель и задачи.

2. Постановка задачи сформулирована не достаточно полно, что является следствием первого замечания. Отсутствует её математическая постановка, учитывающая критерии точности, стабильности и времени классификации, которые имеют вероятностный характер. Подобный подход позволил бы использовать аналитические методы исследований.

3. Автор ограничивается исследованием только метода автоматической классификации типа k -средних и его модификаций, отличающихся используемой метрикой. Поэтому возникает вопрос об обоснованности выбора метода k -средних в качестве основы автоматической классификации промышленной продукции в пространстве их признаков.

4. Остаётся открытой проблема изменения вероятностных характеристик исходных статистических данных при снижении их размерности с применением методов факторного анализа и использовании различных метрик. Нелинейное преобразование случайных величин сопряжено с изменением (вплоть до смещения) количественных характеристик их законов распределения, что приводит к искажению условий автоматической классификации.

5. Не указывается достоверность отличия результатов вычислительных экспериментов при сравнении модификаций алгоритмов автоматической классификации типа k -средних, представленных в таблицах 2.3.2, 2.3.3, 2.3.4.

6. Из 17 публикаций автора только 1 работа подготовлена без соавторов. Возникает необходимость оценить личный вклад автора в других публикациях.

7. В тексте работы допускаются ошибки и некорректные высказывания. Например, на странице 133 в продолжении таблицы 3.9 все столбцы обозначены одним и тем же символом Min; на странице 8 «Методологической базой являются работы по методам кластеризации и классификации»; на странице 11 «... выборка объектов исследования $S=\{0^{(1)}, \dots, 0^{(N)}\}$ ». Не определён индекс i в формуле (1) автореферата.

Следует отметить, что указанные недостатки не носят принципиального характера и не снижают ценность работы.

Заключение

Диссертационная работа Шкабериной Гузели Шарипжановны «Модели и алгоритмы автоматической классификации продукции» имеет внутреннее единство и является завершённой научно-квалификационной работой, в которой содержится решение задачи автоматической группировки продукции, имеющей значение для развития новых методов, обеспечивающих повышение точности и стабильности результатов решения задач автоматической группировки объектов. Выводы, сформулированные в диссертации, обоснованы, обладают научной новизной и имеют практическую значимость.

Автореферат соответствует тексту диссертации и даёт полное представление о содержании и результатах диссертационной работы.

Диссертация соответствует заявленной специальности.

Диссертация соответствует критериям, установленным пунктами 9-11, 13, 14 «Положения о присуждении ученых степеней», утв. постановлением Правительства Российской Федерации от 24.09.2013 № 842 «О порядке присуждения ученых степеней», а ее автор Шкаберина Гузель Шарипжановна достойна присуждения ученой степени кандидата технических наук по специальности 05.13.01– Системный анализ, управление и обработка информации (космические и информационные технологии).

Диссертационная работа была обсуждена и получила положительную оценку на заседании научного семинара отдела прикладной информатики Института вычислительного моделирования Сибирского отделения Российской академии наук – обособленного подразделения ФИЦ КНЦ СО РАН, протокол № 1 от 12.01.2021 г.

Отзыв составила д.т.н., профессор Ноженкова Людмила Федоровна – заведующий отделом прикладной информатики Института вычислительного моделирования СО РАН – обособленного подразделения ФИЦ КНЦ СО РАН.

Федеральное государственное бюджетное научное учреждение
«Федеральный исследовательский центр «Красноярский научный центр
Сибирского отделения Российской академии наук» (ФИЦ КНЦ СО РАН).

Web-сайт организации: <http://icm.krasn.ru/>

660036, г. Красноярск, Академгородок, д. 50, стр. 44.

Телефон: +7 (391) 243 27 56.

Адрес электронной почты: sek@icm.krasn.ru

Заведующий отделом прикладной информатики Института вычислительного моделирования Сибирского отделения Российской академии наук – обособленного подразделения ФИЦ КНЦ СО РАН,
д.т.н., профессор

Л.Ф. Ноженкова

«22» 01 2021 г.

Подпись Л.Ф. Ноженковой заверяю.

Ученый секретарь ИВМ СО РАН

к.ф.-м.н.



А.В. Вяткин

«22» 01 2021 г.