

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА
на диссертационную работу Адуенко Александра Александровича
«Выбор мультимodelей в задачах классификации»,
представленную на соискание ученой степени
кандидата физико-математических наук по специальности
**01.01.09 – «Дискретная математика и
математическая кибернетика».**

Актуальность работы. Диссертационная работа А.А. Адуенко посвящена исследованию актуальной фундаментальной проблемы выбора мультимodelей, содержащих попарно различные модели, в задачах классификации. Данная проблема имеет непосредственную связь с областью информационного поиска, байесовской статистикой и теорией распознавания. Ее исследованию посвящены многочисленные работы отечественных и зарубежных ученых (В.И. Малюгин, В.В. Стрижов, Д. МакКэй, А. Гелман, К. Бишоп и др.)

В диссертационной работе А.А. Адуенко разрабатываются методы классификации, позволяющие учитывать неоднородность данных путем построения мультимodelей, содержащей несколько одиночных моделей, при этом эти модели должны быть выбраны так, чтобы они в каком-то смысле не были близки друг к другу. Для построения адекватных мультимodelей, содержащих попарно различные модели, в работе А.А. Адуенко предлагается функция сходства апостериорных распределений параметров моделей, определенная в случае несовпадения носителей, и изучены ее свойства. На основании свойств распределения функции сходства для пары совпадающих моделей предложены методы выбора адекватных мультимodelей. Рассмотрена и решена задача построения оптимальной адекватной мультимodelей на основании принципа максимума обоснованности и показана асимптотическая вырожденность оценки максимума обоснованности для ковариационной матрицы параметров логистической модели. Актуальность проведенного в диссертации исследования, посвященного методам выбора мультимodelей в задачах классификации, представляется бесспорной.

Научная новизна проведенных исследований и полученных результатов. В диссертации на основе статистического подхода к выбору адекватных мультимodelей и функция сходства апостериорных

распределений параметров для решения задачи сравнения моделей построена теория синтеза адекватных мультимodelей, составленных из попарно статистически различимых моделей. Разработан метод совместного обучения и оптимизации смеси моделей, а также методы комбинирования признаков. Эти результаты и сам предлагаемый подход являются новыми.

Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации.

Научные положения, выводы и рекомендации, сформулированные в диссертации, являются в достаточной степени обоснованными и подтверждены математическими теоремами. Предложенные автором решения аргументированы, проведена оценка их достоверности сравнением с другими известными решениями.

Форма представления материала работы и полученные выводы достоверными.

Оценка содержания диссертации, её завершенность.

Диссертационная работа состоит из введения, пяти глав и заключения.

Во введении обоснована актуальность работы, определены цели и задачи исследования.

В первой главе введены основные понятия и обозначения. Сформулированы понятия оптимальности и обучения вероятностных моделей классификации. Описан мультимодельный подход к задаче классификации.

Во второй главе рассмотрены принцип максимума обоснованности и основанная на нем задача построения оптимальной мультимодели. Доказано, что недиагональная оценка максимума обоснованности ковариационной матрицы параметров логистической модели является асимптотически вырожденной. Предложена оптимальная схема линейного комбинирования признаков с учетом их мультиколлинеарности.

В третьей главе рассмотрена задача обучения мультимodelей. Предложен алгоритм совместного обучения и оптимизации смесей моделей.

В четвертой главе поставлена задача статистического сравнения моделей и построения адекватных мультимodelей. Предложена функция сходства апостериорных распределений параметров моделей для решения задачи сравнения моделей и исследованы ее свойства. Доказаны теоремы об асимптотическом распределении значений функции сходства для совпадающих обобщенно-линейных моделей. На основании этих результатов предложены методы прореживания мультимodelей для

построения адекватных мультимodelей. Получены верхняя и нижняя оценки на максимальное число попарно различных моделей в мультимodelей.

В пятой главе описан программный комплекс, разработанный для решения задач классификации и комбинирования признаков. Приведен анализ результатов на синтетических и реальных данных и сравнение с известными методами.

Основные результаты, полученные в диссертационной работе А.А. Адуенко.

1. Предложена функция схождения плотностей апостериорных распределений параметров моделей, удовлетворяющая требованиям к функции схождения для решения задачи сравнения моделей.
2. Получены верхняя и нижняя оценки на максимальное число моделей в адекватной мультимodelей.
3. Разработаны методы выбора адекватных оптимальных обученных мультимodelей, содержащих попарно статистически различные модели. Предложен метод совместного обучения и отбора признаков для смеси моделей.
4. Предложен метод комбинирования мультиколлинеарных признаков. Доказана асимптотическая вырожденность недиагональной оценки ковариационной матрицы параметров логистической модели, полученной из принципа максимума обоснованности.
5. Разработан программный комплекс для построения адекватных оптимальных обученных мультимodelей при решении задач классификации. Проведены вычислительные эксперименты, подтверждающие адекватность методов.

Таким образом, выполненная работа является целостной, завершенной, имеет теоретическую и практическую значимость

Замечания по диссертационной работе в целом.

1. Концовка раздела 2.1 выглядит скомканной, ее стоило оформить в виде отдельного утверждения.
2. Изложение алгоритмов выбора адекватных мультимodelей в разделе 4.6 можно было бы произвести совместно для смесей моделей и многоуровневых моделей в силу их похожести.

Отмеченные недостатки не носят принципиального характера и не влияют на главные теоретические и практические результаты диссертации.

Результаты диссертационной работы изложены в четырнадцати статьях, из которых девять входят в перечень ВАК РФ. Результаты докладывались на российских и международных конференциях по анализу данных и исследованию операций.

Автореферат адекватно отражает содержание диссертационной работы. **Диссертация соответствует критериям, установленным Положением о присуждении ученых степеней.** Диссертация обладает внутренним единством, содержит новые научные результаты и рекомендации по их практическому использованию.

Работа полностью соответствует требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук по специальности 01.01.09 – Дискретная математика и математическая кибернетика – а ее автор, **А.А. Адуенко, заслуживает присуждения ему ученой степени кандидата наук по данной специальности.**

Официальный оппонент

Доктор физико-математических наук, профессор кафедры математического моделирования и информатики физического факультета Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова»

Адрес: 119991, Российская Федерация, Москва, ГСП-1, Ленинские горы

Тел.: 8- 495-939-41-78, 8-903-570-75-50.

E-mail: achulichkov@gmail.com

3 марта 2017 года

Чу
Чуличков Алексей Иванович

Подпись д.-ф.м.н., проф. Чуличкова А.И. удостоверяю.

Декан физического факультета

МГУ им. М.В.Ломоносова,

профессор



Н.Н.Сысоев
Н.Н.Сысоев