

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский университет ИТМО»
(Университет ИТМО)

Кронверкский пр-т, д. 49, лит. А,
Санкт-Петербург, Россия, 197101
Тел.: (812) 480-00-00 | Факс: (812) 232-23-07
od@itmo.ru | itmo.ru

10.11.2022 № 88-59-01/333

Утверждаю
Проректор по научной работе
университета ИТМО

д.т.н., профессор
Никифоров Владимир Олегович

ОТЗЫВ

ведущей организации – Федерального государственного
автономного образовательного учреждения высшего образования
«Национальный исследовательский университет ИТМО» (Университет ИТМО)
на диссертационную работу Тороповой Александры Витальевны
на тему «**Методы и алгоритмы обработки неопределенности данных при оценивании
интенсивности пуассоновского процесса по ограниченному объему доступных наблюдений**»,
представленную на соискание ученой степени
кандидата технических наук по специальности

2.3.1 – «Системный анализ, управление и обработка информации, статистика»

Актуальность темы

Диссертационная работа А. В. Тороповой посвящена решению актуальной научной проблемы — обработки неопределенности данных респондентов в задаче оценивания интенсивности пуассоновского процесса, выступающего моделью эпизодического поведения. Исследования поведения проводятся во многих областях науки: в социологии, экономике, медицине и др. В связи с тем, что в большинстве случаев прямое наблюдение за поведением индивида является невозможным по различным причинам, основным источником данных о поведении становятся результаты опросов или анкетирования респондентов. Данные, собранные таким образом, могут быть неточны и ограничены по объему. Вследствие этого требуется обработка неопределенности таких данных для получения более точных оценок параметров изучаемого поведения. В диссертационной работе рассматриваются подходы к обработке неопределенности данных респондентов, возникающей в задаче оценивания одного из ключевых параметров эпизодического поведения — его интенсивности. В условиях ограниченности ресурсов результаты диссертационной работы могут быть востребованы специалистами, изучающими поведение человека, так как на основе небольшого количества данных с помощью предложенных методов и алгоритмов можно получить довольно высокое качество оценивания интенсивности эпизодического поведения.

Содержание

Диссертационная работа состоит из введения, четырех глав, заключения и приложений. Во введении приведена актуальность работы, установлены цели и задачи, перечислены основные выносимые на защиту положения, приведен краткий обзор методов исследования, установлена научная новизна, теоретическая и практическая значимость, приведена достоверность результатов, перечислены основные публикации по теме диссертации.

В первой главе обоснованы актуальность диссертационного исследования, выбор байесовских сетей доверия как основного инструмента. Проведен обзор сфер применения байесовских сетей доверия, инструментария для работы с ними, возможные подходы к их

обучению.

Вторая глава посвящена описанию теоретических результатов, послуживших предпосылками для диссертационного исследования. Сделана формальная постановка задачи, описана схема диссертационного исследования.

Третья глава описывает полученные соискателем теоретические результаты. Предложены метод и алгоритм оценивания согласованности информации о последних эпизодах и рекордных интервалах пуассоновского процесса в модели оценивания его интенсивности, основанной на байесовской сети доверия; алгоритм обработки некорректности информации, полученной от респондентов, при оценивании интенсивности пуассоновского процесса, выступающего математической моделью поведения индивида; метод и алгоритм обработки длины интервала между последним эпизодом процесса и концом исследуемого периода, при оценивании интенсивности пуассоновского процесса по ограниченному объему доступных наблюдений. Рассмотрена модель оценивания интенсивности пуассоновского процесса, расширенная скрытыми переменными, как истинными данными о последних эпизодах и рекордных интервалах процесса, со структурой, обученной на синтетических данных. Описаны модели оценивания интенсивности постинга в социальных сетях, расширенные за счет объективных данных о пользователе. Рассмотрены возможные варианты дискретизации непрерывных величин, входящих в модели.

В четвертой главе описаны архитектура и прототип комплекса программ для работы с моделями оценивания интенсивности пуассоновского процесса, реализующего разработанные методы и алгоритмы. Описаны использованные для апробации предложенных алгоритмов данные, синтетические и данные из социальных сетей, и результаты этой апробации. Описаны внедрения полученных результатов.

Основные результаты и их новизна

В рамках диссертационной работы А. В. Тороповой получены следующие новые результаты:

1. Предложены новые метод и алгоритм оценивания согласованности информации о последних эпизодах и рекордных интервалах пуассоновского процесса, основанные на расширении байесовской сети доверия, ранее применяемой для оценивания интенсивности пуассоновского процесса, дополнительными узлами, характеризующими согласованность данных респондентов, что может обеспечить возможность работы с данными необходимой степени согласованности.

2. Разработан новый алгоритм обработки некорректности данных респондентов при оценивании интенсивности пуассоновского процесса за счет использования скрытых переменных в байесовской сети доверия, отвечающих истинным значениям об эпизодах поведения респондентов. По сравнению с предложенным ранее подходом была повышена точность оценки интенсивности пуассоновского процесса.

3. Разработан новый алгоритм обработки неопределенности задания конца исследуемого периода в задаче оценивания интенсивности пуассоновского процесса по ограниченному объему доступных наблюдений. Алгоритм основан на введении в байесовскую сеть доверия вершины, характеризующей интервал между последним эпизодом пуассоновского процесса и эпизодом, произошедшим после окончания периода исследования. По сравнению с предложенным ранее подходом была повышена точность оценки интенсивности пуассоновского процесса.

4. Разработана программная реализация предложенных в диссертационном исследовании новых алгоритмов обработки неопределенности данных при оценивании интенсивности пуассоновского процесса по ограниченному объему доступных наблюдений.

Результаты могут быть применены в научных исследованиях и прикладных задачах, в частности в сфере управления персоналом, в которых важно оценивать числовые характеристики поведения человека. Сказанное позволяет рекомендовать применение результатов в различных научных и коммерческих организациях, например, ФГАОУ ВО НИУ ИТМО, ФГБУН СПб ФИЦ РАН, ФГБОУ ВО «РАНХиГС» и др.

Достоверность результатов

Обоснованность и достоверность научных положений, представленных в диссертационной работе, обеспечивается за счет качественного анализа опубликованных исследований по тематике, корректности постановки задачи исследования, согласованности полученных результатов, обсуждением основных результатов на международных и российских научных конференциях, успешными внедрениями и поддержкой грантами РФФИ. Также можно отметить, что соискателем

опубликованы 42 научные и приравненные к ним работы, в частности имеются 4 публикации в изданиях из перечня рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, 11 публикаций в изданиях, индексируемых Scopus/WoS.

Значимость результатов для развития соответствующей отрасли науки

Основными научными результатами автора является разработка новых методов, моделей и алгоритмов оценивания интенсивности пуассоновского процесса, выступающего математической моделью эпизодического поведения индивида, обрабатывающие некоторые типы неопределенности данных респондентов.

Полученные результаты создают основу для обработки неопределенности данных респондентов, при построении оценок интенсивности эпизодического поведения человека по ограниченному объему доступных наблюдений. Также предложенные методы и алгоритмы обеспечивают возможность автоматизации решения задачи построения оценок и составляют основу для создания систем поддержки принятия решений в областях знаний, изучающих поведение человека.

Результаты диссертационного исследования получили практическую реализацию в ООО «Хоум Фитнес» при разработке подходящего для клиента режима физических нагрузок, были внедрены в научно-исследовательскую работу ФГБУН СПб ФИЦ РАН, в которой решалась задача автоматизации сбора информации об эпизодическом поведении индивида и вычислении сводных характеристик такого поведения, а также использованы в учебном процессе СЗИУ РАНХиГС. Внедрения подтверждены соответствующими актами.

Замечания

1. В главе 1 сказано: «В диссертационном исследовании решается задача обработки некоторых типов неопределенности данных при оценивании интенсивности пуассоновского процесса по ограниченному объему доступных наблюдений». Следует указать, какие еще могут возникнуть типы неопределенности данных в этой задаче, и почему они не рассматриваются в работе.

2. Для опроса респондентов о поведении также часто используется календарный метод, при его использовании респондентам надо отметить в календаре, в какие дни происходили эпизоды исследуемого поведения, в работе можно было рассмотреть и такой метод для сбора информации.

3. Разделы 1.3–1.5 и 2.1 посвящены байесовским сетям доверия, однако выбор этого инструмента обоснован достаточно кратко, возможно, стоило бы привести примеры применения других инструментов для решения исследуемой задачи.

4. В работе следовало бы обосновать выбор моделей ошибки для доказательства повышения качества оценки классификации в разделе «4.6 Внедрение методов и алгоритмов обработки неопределенности при оценивании интенсивности пуассоновского процесса», также если рассматривались другие модели ошибки, следовало бы также описать полученные результаты.

5. В разделе 3.1 обнаружено противоречие: «Вероятность согласованности означиваний случайных величин $\langle \dots \rangle$, которые соответствуют интервалам между последним и предпоследним эпизодами пуассоновского процесса $\langle \dots \rangle$ рассчитывается следующим образом: если данные противоречат друг другу (интервал меньше минимального), то она равна 0, если данные согласованы — 1», однако из дальнейших объяснений следует, что за счет различных значений порогов согласованности, у этой величины возможны также и значения отличные от 0 и 1.

6. В работе в качестве примера применения полученных результатов рассматривается эпизодическое поведение, имело бы смысл указать, подходят ли предложенные методы и алгоритмы для любого эпизодического поведения, или существуют какие-то ограничения.

7. Не представлена информация о том, как собиралась выборка респондентов, заполнивших разработанный автором опросный инструментарий.

8. Интерфейсы разработанных программ (рис. 4.3–4.29) представлены на английском языке, предусмотрена ли локализация для русскоязычных пользователей? Если нет, то следует проработать этот вопрос.

9. В работе имеют место недостатки методического характера: присутствуют ошибки в оформлении работы, в частности у таблицы 3.5 заголовок расположен на одной странице, а сама таблица – на следующей, в таблице 3.6 содержимое всех колонок выровнено по центру, кроме

последней, в которой содержимое выровнено по левому краю, и др.; также встречаются пунктуационные ошибки, в частности, на с. 5 лишняя запятая перед «и др.» в первом абзаце, в сложноподчиненном предложении на с. 19, предпоследнем во втором абзаце пропущена запятая, и др.

Заключительная оценка

Диссертационная работа Тороповой Александры Витальевны «Методы и алгоритмы обработки неопределенности данных при оценивании интенсивности пуассоновского процесса по ограниченному объему доступных наблюдений» является законченной научно-квалификационной работой, содержащей решение актуальной научной задачи – обработки неопределенности данных при оценивании интенсивности эпизодического поведения, моделью которого выступает пуассоновский процесс. Представленные в работе результаты и выводы обоснованы корректно. Результаты работы своевременно и полно опубликованы в 42 печатных трудах, 4 из которых – в изданиях, рекомендованных ВАК РФ, 11 — в изданиях, индексируемых Scopus/WoS. Результаты докладывались автором на российских и международных конференциях. Автореферат полностью отражает основное содержание диссертации.

Диссертация соответствует всем критериям, установленным п. 9–14 «Положения о присуждении ученых степеней» утвержденным Постановлением Правительства РФ №842 от 24 сентября 2013 «О порядке присуждения ученых степеней», предъявляемым к диссертациям на соискание ученой степени кандидата технических наук по специальности 2.3.1 – «Системный анализ, управление и обработка информации, статистика», а ее автор – Торопова Александра Витальевна – заслуживает присуждения ученой степени кандидата технических наук по указанной специальности.

Настоящий отзыв обсужден и одобрен на расширенном научном семинаре факультета информационных технологий и программирования Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет ИТМО» (протокол № 11 от 8 ноября 2022)

д.т.н., профессор ИТиП

Шалыто АА

к.т.н., доцент ИТиП

Забашта А.С.

Почтовый адрес: 197101, г. Санкт-Петербург, Кронверкский пр-т, д. 49, лит. А,

Тел.: (812) 480-00-00 | Факс: (812) 232-23-07

|Адрес электронной почты: od@itmo.ru

Организация – место работы: федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО» (Университет ИТМО)

Web-сайт организации: itmo.ru