

Отзыв официального оппонента

доктора технических наук, профессора, Иванова Александра Ивановича, научного консультанта АО «Пензенский научно-исследовательский электротехнический институт» на диссертационную работу Зоткиной Алены Александровны «Методы и алгоритмы формирования психологического портрета пользователя социальной сети для эффективного подбора кадров», представленную на соискание ученой степени кандидата технических наук по специальности 2.3.8. – «Информатика и информационные процессы» (отрасль науки – технические) в объединенный диссертационный совет 99.2.113.02 при ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», ФГБОУ ВО «Пензенский государственный технологический университет»

Актуальность темы исследования

В современном обществе социальные сети генерируют огромные объемы данных, которые можно использовать для анализа личностных качеств и настроений человека, формируя его портрет. С увеличением объема, анализируемой информации, возникает необходимость в разработке новых методов анализа, позволяющих оценивать эмоциональные изменения портрета личности, включая оценку уровня депрессивных состояний. Применение статистических методов и нейронных сетей для обработки данных открывает возможности для более точного прогнозирования поведения проверяемых и оптимизации оказания им психологической поддержки в случае необходимости.

Многие текущие исследования сосредотачиваются на анализе данных лишь в одной социальной сети, что снижает точность предсказаний из-за ограниченного объема исходной информации. Интеграция данных из разных источников и учет аспектов пользовательской активности необходимы для создания более точных и более сложных психологических портретов (растет входная размерность решаемой задачи).

Таким образом, тема диссертационного исследования является актуальной и имеет значение как для совершенствования научной методологии, так и для практического применения при рекрутинге. Работодатели всё чаще стремятся минимизировать риски, связанные с

процессом найма, опираясь на научное обоснование, а также технические решения, способные помочь в формировании более точных психологических портретов. Новые технологии значительно повышают как безопасность, так и эффективность управления человеческими ресурсами.

Обоснованность и достоверность научных положений, выводов и рекомендаций.

Обоснованность научных положений, выводов и результатов, приведенных в диссертационной работе Зоткиной А.А., определяются корректным применением современных методов исследования: интеллектуального анализа данных, теории вероятностей и математической статистики для обработки экспериментальных данных, методов обработки естественного языка, методами теории анализа социальных сетей (Social Network Analysis, SNA).

Достоверность полученных результатов подтверждается результатами проведенных экспериментов, а также апробированием основных полученных результатов исследования на Международных и Всероссийских научных конференциях.

Результаты диссертации отражены в 38 работах, в том числе в 6 статьях в журналах, рекомендованных ВАК Минобрауки России, в 2 статьях, индексируемых в международной наукометрической базе Scopus, и в 2 свидетельствах о государственной регистрации программ для ЭВМ.

Общий анализ содержания работы.

Диссертационная работа изложена на 140 страницах машинописного текста, и состоит из списка сокращений, введения, четырех разделов, основных результатов и выводов, списка литературы из 133 наименований, включающего работы отечественных и зарубежных авторов, и приложений. Диссертация содержит 29 рисунков и 9 таблиц. Материалы, изложенные в автореферате, достаточно полно отражают содержание диссертации.

Во введении представлены актуальность темы диссертации, ее цель и задачи, обозначена постановка существующей проблемы, описано текущее состояние вопроса по рассматриваемой области исследования, кратко сформулированы и обоснованы результаты научного исследования.

В первой главе проводится аналитический обзор основных методов интеллектуального анализа текстовых данных в социальных сетях, выявляются ключевые проблемы, сопровождающие процесс построения психологического портрета пользователей. Анализируются существующие подходы к обработке данных из социальных платформ, акцентируя внимание на ограничениях, связанных с использованием информации лишь из одной социальной сети, что ведет к снижению предсказательной способности моделей. Также рассматриваются современные методы оценки личностных черт, выделяются их достоинства и недостатки. Формулируется обоснованный вывод о необходимости интеграции данных с различных платформ для более точного анализа пользователей и улучшения качества построения психологических профилей (портретов). На основе проведенного аналитического обзора определяются ключевые задачи диссертационного исследования.

Во второй главе проводится детальное описание метода для сравнения характеристик выражений и текстовых сообщений пользователей с аналогичными аккаунтами в социальных сетях. Отдельно приводится описание метода интеграции данных, размещаемой пользователем в разных социальных сетях, который позволяет восстанавливать данные активности, учитывая разнообразные аспекты его онлайн-поведения, для составления более полного и подробного психологического портрета и определения отклоняющегося поведения. Приводится описание методики кросс-доменного аспектно-ориентированного анализа тональности текста IbDA-LSTM-CRF, которая решает проблему аспектно-ориентированного анализа тональности. Предложенная методика учитывает как общий контекст, так и особенности каждого текста, вне зависимости от его тематики или

смыслового контекста. Это позволяет проводить более глубокий анализ тональности и оценивать эмоциональные аспекты текста в различных предметных областях, что повышает универсальность и применимость моделей для работы с различными категориями данных.

Рассматриваемые в работе функционалы строятся на сравнении подобия текстов и их фрагментов с приведением каждого из параметров к интервальной арифметике $[0, 1]$. Обобщение множества параметров выполнено через взвешенное суммирование. Анализируемые текстовые сообщения разбиваются на большое число фрагментов, выполняется автоматическая BIO-разметка данных для последующего их нейросетевого анализа.

В третьей главе описан разработанный программный комплекс для формирования психологического портрета пользователя социальной сети. Программа разработана на основе предложенных автором методов, описанных во втором разделе. Алгоритмы, представленные в диссертационной работе и автореферате, ориентированы на вероятностные оценки, выполняемые тремя нейросетевыми моделями.

В четвертом разделе приводится описание экспериментальной части исследования. В ходе исследования был проведен сравнительный анализ профилей пользователей во «ВКонтакте» и «Одноклассниках» на выборке из 153 человек (261 аккаунт), что выявило критерии точной идентификации учетных записей. Также были проведены эксперименты по формированию психологического профиля и вероятности отклоняющегося поведения с помощью нейронной сети LSTM, обученной на 29000 записях активности пользователей. Модель продемонстрировала достаточно высокую точность, превосходя другие алгоритмы (KNN и Random Forest). Доверительная вероятность к нейросетевой классификации достигает 0,93-0,96. Это свидетельствует об адекватности алгоритмов обработки результатов.

Приводится таблица сравнения доверительных вероятностей для трех типов нейронных сетей в ситуации разделения пользователей на 16 классов

психологических портретов. Показано, что нейросеть с архитектурой LSTM дает наиболее достоверные оценки.

Приводятся результаты численных экспериментов для трех сравниваемых алгоритмов (сеть - LSTM, KNN-соседи, Random Forest) в терминах доверительной вероятности (точности).

Научная новизна диссертации заключается в достижении следующих обоснованных и достоверных результатов:

1. Разработан метод оценки сходств признаков выражения, текстовых объектов, записей множества пользователей социальных сетей и реализующий ее алгоритм работы поиска аккаунтов пользователя, которые в отличие от существующих, учитывают разнообразные аспекты активности пользователей (публикации, участие в сообществах, комментарии, лайки к комментариям и публикациям). Это позволяет более точно идентифицировать одинаковые аккаунты пользователей.

2. Разработан метод интеграции информации, публикуемой пользователем на разных платформах социальных сетей, позволяющий восстанавливать данные о пользователях, проявивших активность хотя бы на одной из этих платформ, который отличается тем, что учитываются полные данные об активности пользователя на протяжении длительного периода времени, что позволяет составить более полный и подробный психологический портрет пользователя.

3. Разработана методика кросс-доменного аспектно-ориентированного анализа тональности текста и алгоритм на ее основе, которая в отличие от существующих, фокусируется на выделении аспектов и анализе тональности отношения к ним в тексте, что позволяет получить более детальное представление о содержании и оценке текста, в отличие от других рассматриваемых методик.

4. Нейросетевая методика и алгоритм, ее реализующий, для определения психологических характеристик пользователя социальной сети, с использованием типологии MBTI, которая в отличие от других подходов,

фокусируется на изолированных личностных чертах, что позволяет предоставить комплексное представление личности.

Практическая значимость работы заключается в том, что использование методов, методик, алгоритмов и программных решений, разработанных в рамках диссертации, способствует сокращению времени формирования психологического портрета пользователя, что позволяет значительно повысить эффективность управления кадровой системой.

Замечания по диссертационной работе:

1. В автореферате отсутствуют оценки входной размерности решаемой нейросетевой задачи при проведении численных экспериментов (при докладе необходимо указать конкретное число анализируемых нейросетями входных параметров).

2. При сравнительных испытаниях трех типов нейросетевых архитектур необходимо обеспечить тождество исходных данных, тождество извлеченных из них функционалов (параметров), тождество выделяемых классов портретов пользователей (при докладе эта неопределенность должна быть снята соискателем).

3. В автореферате отсутствует подробное описание рисунков 1-3, что вызывает некоторые затруднения при восприятии данного аспекта работы.

4. В автореферате недостаточно сведений о том, каким образом были рассчитаны переменные Precision, Recall и F-мера.

5. В тексте диссертации и автореферата присутствуют стилистические ошибки и опечатки.

Заключение

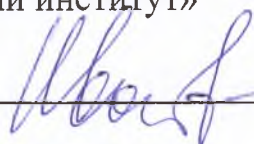
Отмеченные замечания не носят принципиального характера, не влияют на общую положительную оценку диссертационного исследования,

не снижают значимость проведенной работы, не влияют на теоретические и практические результаты диссертационного исследования и не ставят под сомнение достоверность полученных выводов и результатов.

Считаю, что диссертация является законченной научно-квалифицированной работой, соответствующей по актуальности и уровню исполнения требования п. 9 «Положение о порядке присуждения ученых степеней» ВАК Министерства высшего образования и науки РФ, предъявляемых к кандидатским диссертациям, а Зоткина Алена Александровна заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.8. – «Информатика и информационные процессы» (отрасль науки – технические).

Официальный оппонент:

доктор технических наук, профессор,
научный консультант АО «Пензенский
научно-исследовательский электротехнический институт»

 А.И. Иванов

28.11.2024

Подпись Иванова А.И. заверяю
зам. генерального директора – научный директор АО «ПНИЭИ»



О.В. Ефимов

Иванов Александр Иванович

АО «Пензенский научно-исследовательский электротехнический институт»

Почтовый адрес (рабочий): 440026, г. Пенза, ул. Советская, 9.

Телефон (рабочий): 8 (8412) 59-33-10; e-mail: bio.ivan.penza@mail.ru

Докторская диссертация защищена в 2002 году по специальности 05.13.01 – Системный анализ, управление и обработка данных.

Иванов А.И. являлся руководителем работ по созданию 7 национальных стандартов России в рамках деятельности ТК 362 «Техника защиты информации», регламентирующих требования к средствам биометрико-нейросетевой аутентификации с 2006 по 2014 г.г.

С отзовом ознакомлена 29.11.24
Зоткина А.А.
7