

С 2018 г. по настоящее время обучается в аспирантуре Ульяновского государственного технического университета по специальности 05.13.12 «Системы автоматизации проектирования (промышленность)». Справки об обучении в аспирантуре и сдаче кандидатского минимума по экзаменам получил в 2022 г.

Научный руководитель – Мошкин Вадим Сергеевич, к.т.н., доцент кафедры «Информационные системы» федерального государственного бюджетного образовательного учреждения высшего образования «Ульяновский государственный технический университет».

По итогам обсуждения принято следующее заключение:

Автор Андреев И.А. выполнил исследование методов и алгоритмов обработки текстовой информации социальных сетей. Автором точно сформулированы, обозначены и решены задачи исследования, направленные на разработку методов и алгоритмов анализа открытых русскоязычных текстовых данных социальных сетей, реализованных в виде комплекса программ, позволяющих автоматизировать построение социального портрета пользователей социальных сетей.

Автор Андреев И.А. определил следующую цель исследования.

Целью диссертационного исследования является снижение трудозатрат на построение социального портрета пользователей социальных сетей посредством автоматизации и учета дополнительных факторов в процессе анализа открытых русскоязычных текстовых данных.

Цель декомпозирована на ряд задач исследования, сформулированных следующим образом:

- анализ существующих работ по формированию обучающих выборок и сентимент-анализу текстовых постов социальной сети;
- сравнение современных интеллектуальных методов анализа текстовых данных, выявление их возможностей и ограничений в рамках психолингвистического и сентимент-анализа данных постов в социальной сети;
- разработка алгоритма формирования обучающей выборки, состоящей из открытых русскоязычных текстовых ресурсов социальных сетей,

классифицированных по 7-ми эмоциям;

- разработка алгоритма классификации текстовых сообщений социальной сети по классам тональности на основе семантических подходов и машинного обучения;
- разработка подхода к сопоставлению профилей пользователей в разных социальных сетях посредством анализа структурированных и неструктурированных данных анкет, а также социальных графов профилей;
- разработка метода определения психологических характеристик пользователя социальных сетей посредством анализа текстовых сообщений в социальных сетях;
- разработка программной системы психолингвистического и сентимент-анализа открытых текстовых русскоязычных данных профилей пользователей социальных сетей;
- проведение вычислительных экспериментов, позволяющих оценить эффективность предложенных методов и алгоритмов;
- внедрение результатов исследования в практику процесса подбора персонала организаций региона.

Объект исследования – набор открытых русскоязычных текстовых данных, извлекаемых со страниц пользователей социальных сетей.

Предмет исследования – модели и алгоритмы психолингвистического и сентимент-анализа русскоязычных текстовых данных социальных сетей.

Научную новизну автор формулирует следующим образом:

1. Разработан алгоритм формирования обучающей выборки для обучения моделей классификации в задачах сентимент-анализа текстовых данных, отличающийся совместным использованием словарей авторских символов выражения эмоций и ключевых фраз;
2. Предложен подход к сопоставлению профилей пользователей в разных социальных сетях, отличающийся гибридизацией подходов анализа графической информации, структурированных данных анкет, текстовых

данных, а также социальных графов профилей;

3. Разработан метод определения психологических характеристик пользователя социальных сетей, отличающийся гибридизацией алгоритмов обработки естественного языка текстовых данных, машинного обучения и метода «Большой пятерки» ;

4. Предложен алгоритм анализа эмоциональной окраски русскоязычных текстовых данных социальных сетей, отличающийся интеграцией семантических подходов и методов машинного обучения.

Практическая значимость диссертационной работы заключается в разработке программного комплекса, позволяющего упростить подбор персонала в организации посредством разработки социального портрета, полученного путем анализа профилей человека в социальных сетях.

Автор выносит на защиту следующие положения:

1. Разработанный алгоритм формирования обучающей выборки позволяет эффективно решать задачу обучения нейронной сети в процессе сентимент-анализа русскоязычных текстов социальных сетей;

2. Предложенный подход к сопоставлению профилей пользователей в разных социальных сетях реализован в программном комплексе и автоматизирует процесс поиска профилей пользователя в задаче построения социального портрета;

3. Предложенный метод определения психологических характеристик пользователя социальных сетей с применением методов машинного обучения и модели «Большой пятерки» позволяет классифицировать пользователя по пяти основным факторам данной модели;

4. Разработанный алгоритм анализа эмоциональной окраски русскоязычных текстовых данных, отличающийся интеграцией семантических подходов и методов машинного обучения, повышает точность классификации текстов социальных сетей по классам тональности.

Реализованный в рамках результатов исследования программный комплекс был использован в рамках проекта «Интеллектуальная платформа формирования социального портрета соискателя на основании семантико-когнитивного анализа

профилей в социальных сетях», поддержанного Фонда содействия инновациям по программе «Старт-Цифровые технологии» для компании ООО «Центр программной инженерии и аналитики «ФаззиЛаб».

Практическая значимость диссертационной работы подтверждается использованием разработанных алгоритмов и методов при разработке системы интеллектуального поиска и анализа в Интернет-СМИ и социальных сетях в рамках совместного проекта с ФНПЦ АО «НПО «Марс».

Кроме того, разработанные алгоритмы и подходы были применены УОСОО «Федерация бадминтона» в рамках проекта «Парабадминтон: все силы – для победы», поддержанного Фондом Президентских грантов для отбора волонтеров, обеспечивающих сопровождение лиц с ПОДА (проект № 18-2-009220).

Апробация результатов исследования.

Основные положения и результаты диссертационной работы опубликованы в журналах из Перечня, рекомендованного ВАК РФ: «Онтология проектирования», «Автоматизация процессов управления», докладывались, обсуждались и получили одобрение на следующих конференциях, семинарах и симпозиумах XXIX Международной конференции «Computational Science and Its Applications»-ICCSA-2019 (г.Санкт-Петербург, 2019 г.); Международной научно-технической конференции «Автоматизация» – RusAutoConf-2020 (г.Сочи, 2020 г.); XII Международной конференции Developments in eSystems Engineering – DESE- 2019 (г.Казань, 2019 г.); V Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (г. Минск, 2015 г.); VIII и IX Международных научно-практических конференциях «Интегрированные модели и мягкие вычисления в искусственном интеллекте» (г. Коломна, 2015, 2020 гг.); I Международной научной конференции «Интеллектуальные информационные технологии в технике и на производстве» (г. Сочи, 2016 г.); III, V, VI, VII Международных

конференциях и молодежных школах «Информационные технологии и нанотехнологии» (г. Самара, 2017, 2019, 2020, 2021); VIII Международной конференции «Системный анализ и информационные технологии» САИТ – 2019 (г. Иркутск, 2019 г.); I Международной Поспеловской летней школе-семинаре для студентов, магистрантов и аспирантов «Методы и технологии гибридного и синергетического искусственного интеллекта» (г. Светлогорск, 2014 г.); V Всероссийской Поспеловской конференции с международным участием «Гибридные и синергетические интеллектуальные системы» (г. Светлогорск, 2020 г.); XVII национальной конференции по искусственному интеллекту с международным участием «КИИ-2019» (г. Ульяновск, 2019 г.); IV Всероссийской научно-практической мультikonференции с международным участием «Прикладные информационные системы»-ПИС-2017 (г. Ульяновск, 2017 г.); 6-й Всероссийской научно-технической конференции аспирантов, студентов и молодых ученых ИВТ-2014 (г. Ульяновск, 2014 г.).

По результатам работы было опубликовано 32 статьи, 4 из которых в журналах из перечня ВАК, 11 статей в изданиях, индексируемых в Scopus и/или Web Of Science, а также 1 монография. Получены 3 свидетельства о государственной регистрации программ для ЭВМ.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы и приложений. Текст работы изложен на 166 страницах, включает в себя 50 рисунков, 10 таблиц. Список использованной литературы – 170 наименований.

Основные результаты, полученные в диссертационном исследовании, соответствуют следующим пунктам специальности паспорта специальности 05.13.01 – «Системный анализ, управление и обработка информации (информационные технологии и промышленность)»:

- п. 4 – разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации;
- п. 10 – методы и алгоритмы интеллектуальной поддержки при принятии управленческих решений в технических, экономических, биологических, медицинских и социальных системах.

Личный вклад соискателя. Представленные в данной работе результаты получены автором самостоятельно. Подготовка к публикации некоторых результатов проводилась совместно с соавторами, причем вклад диссертанта был определяющим.

Исходя из вышеизложенного, следует считать, что работа Андреева Ильи Алексеевича выполнена на важную научно-техническую тему, является актуальной, представляет собой законченное научное исследование и решает поставленные задачи в полном объеме.

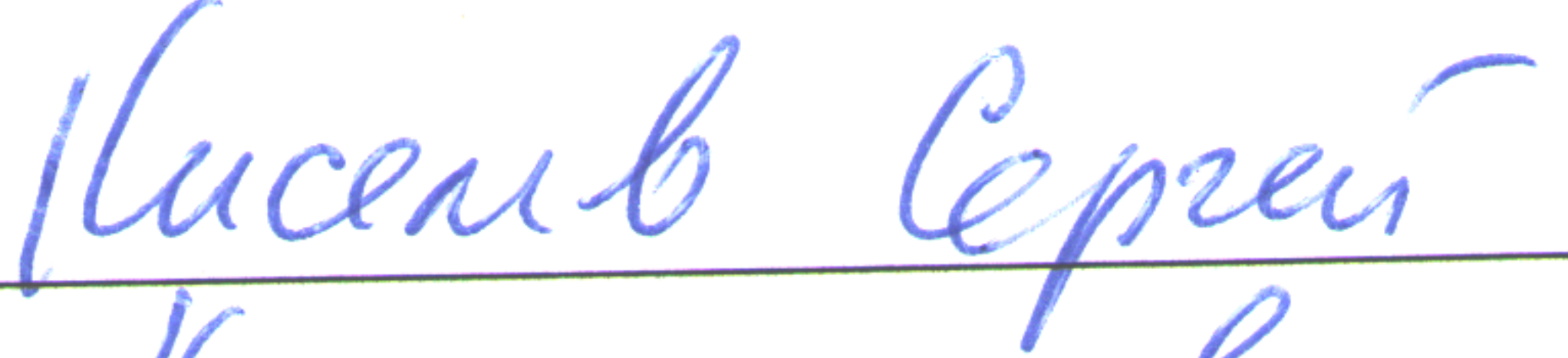
Диссертация Андреева Ильи Алексеевича «Исследование методов и алгоритмов обработки текстовой информации социальных сетей в задачах формирования социального портрета пользователя» рекомендуется к защите на соискание ученой степени кандидата технических наук по специальности 05.13.01 – «Системный анализ, управление и обработка информации (информационные технологии и промышленность)».

Заключение принято на заседании Научно-технического совета факультета информационных систем и технологий (ФИСТ) УлГТУ.

Присутствовали на заседании 11 чел. Результаты голосования: «за» - 11 чел., «против» - 0 чел., «воздержались» - 0 чел., протокол № 5 от «12» мая 2022 г.



(подпись лица оформившего заключение)



Константинов,

Д.т.н., доцент,

заведующий ка-

федрой Информатика-

вычислительные компи-

лекей", председатель

НТС, ФИСТ

(фамилия, имя, отчество – при наличии, ученая степень,
ученое звание, наименование структурного подразделения,
должность)