

ОТЗЫВ

Члена диссертационного совета доктора технических наук
Дегтярева Александра Борисовича
на диссертацию Тарасова Никиты Андреевича на тему
«Гибридные нейросетевые методы анализа понятности текстов юридических документов на русском языке», представленную на соискание ученой степени кандидата технических наук по научной специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Общая информация о диссертации

Диссертация Тарасова Н. А. состоит из введения, пяти глав, заключения, списка литературы, включающего 156 наименований использованных источников. Общий объем текста на русском языке составляет 124 страницы, включая 21 рисунок и 14 таблиц.

Материалы диссертации опубликованы в 9 печатных работах, из них 4 статьи в изданиях Scopus и WoS, 5 в тезисах докладов, получены 3 свидетельства о государственной регистрации программ для ЭВМ.

Актуальность темы исследования

Автоматизированные методы обработки текстов позволяют решать широкий спектр задач, связанных с анализом юридической документации. Интеллектуальные методы анализа данных позволяют структурировать, систематизировать и классифицировать документы по различным языковым характеристикам, в том числе по сложности. Определение сложности и доступности правовых текстов способно помочь улучшить понятность официальных документов, способствуя лучшему взаимопониманию между юристами и лицами, не обладающими профессиональными юридическими знаниями.

Достоверность выводов и результатов исследования

Достоверность выполненных автором исследований подтверждена результатами вычислительных экспериментов и адекватностью использованных в работе разработанных автором методов и алгоритмов.

Научная новизна исследований и полученных результатов

Проведенное исследование вносит весомый вклад в изучение сложности и доступности юридического языка. Разработка нового гибридного подхода, интегрирующего нейросетевые модели и языковые характеристики позволила эффективно выявлять и анализировать особенности юридических текстов, оказывающих наибольшее влияние на сложность их восприятия.

Теоретическая и практическая значимость работы

Исследование и его результаты имеют важное теоретическое и практическое значение. Разработан комплекс методологических подходов и алгоритмов, а также программные средства, способные давать оценку сложности восприятия юридических документов. Посредством статистических методов и методов машинного обучения были определены характеристики, которые наиболее эффективно описывают сложность юридических текстов. В сочетании с нейросетевыми языковыми моделями был создан

программный инструмент для определения сложности правовых текстов. Разработанные методы анализа имеют широкий круг применимости, в частности решают один из неотъемлемых вопросов улучшения правовой осведомлённости.

Анализ содержания работы

Во введении содержится следующая информация: актуальность темы исследования; цель и задачи; научная новизна; теоретическая и практическая значимость работы; апробация результатов; информация о публикациях; структуре и объеме работы; основные научные результаты; положения, выносимые на защиту. В первой главе анализируются методы оценки частотности употребления юридических терминов в различных типах документов с применением статистических методов. Во второй главе исследуются признаки, определяющие понятность юридических документов. Проводится анализ эффективности этих признаков применительно к задаче классификации документов по уровню сложности. В третьей главе представлен гибридный метод оценки сложности, объединяющий языковые характеристики и нейросетевые языковые модели. В четвертой главе осуществляется сравнительный анализ сложности юридических документов различных типов и стилей с использованием гибридной семантической модели. Описана методология сбора, предварительной обработки и анализа документов из различных источников. В пятой главе демонстрируется практическое применение разработанных методов на текстах вопросов и ответов на юридические запросы в области налогового права.

При анализе текста диссертации появились следующие замечания:

1. В работе рассмотрено 130 различных метрик, которые сопоставляются с понятием «простой/средний/сложный» текст. При этом четкой формализации последнего нет. Утверждение, что сложность текста можно соотнести с учебниками для разных классов школы кажется спорным. Для этих целей необходимо было ввести интегральную метрику. Кроме того, сложность юридического текста и просто сложность текста по школьным учебникам разные вещи.
2. На многих рисунках присутствует большое количество фраз на английском языке (например, 4.3-4.6), хотя все эти обозначения относятся к русским терминам, например, RF – это, видимо, РФ (Российская Федерация). Понятно, что взято из компьютерных программ, но в тексте диссертации это приводит к ухудшению понимания содержания.
3. В работе имеются опечатки, лишние знаки препинания, дублированные слова, в, в принципе, в большинстве текстов диссертаций.

Заключение

Диссертация представляет собой целостную и законченную научно-исследовательскую работу на актуальную тему, выполненную автором самостоятельно на хорошем научно-техническом уровне. Основные выводы по результатам исследований достоверны и обоснованы. Опубликованные научные работы подробно отражают основное содержание диссертации и полученные выводы.

Диссертационная работа Тарасова Никиты Андреевича «Гибридные нейросетевые методы анализа понятности текстов юридических документов на русском языке» по теоретическому уровню и практической значимости соответствуют требованиям «Положения о присуждении ученых степеней», предъявляемым к кандидатским диссертациям, а её автор, Тарасов Никита Андреевич, заслуживает присуждения ученой

степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Диссертация Тарасова Никиты Андреевича на тему: «Гибридные нейросетевые методы анализа понятности текстов юридических документов на русском языке» соответствует основным требованиям, установленным Приказом от 19.11.2021 № 11181/1 «О порядке присуждения ученых степеней в Санкт-Петербургском государственном университете», соискатель Тарасов Никита Андреевич заслуживает/ присуждения ученой степени кандидата технических наук по научной специальности 2.3.1. Системный анализ, управление и обработка информации, статистика. Нарушения пунктов 9 и 11 указанного Порядка в диссертации не обнаружены.

Член диссертационного совета

Доктор технических наук, доцент

Профессор кафедры компьютерного моделирования и многопроцессорных систем СПбГУ,
19804, Санкт-Петербург, Университетский пр., 35,

E-mail: a.degtyarev@spbu.ru

Дегтярев Александр Борисович



Дата 05.12.2024

Подпись