

## **ОТЗЫВ**

Научного руководителя, кандидата технических наук Блеканова Ивана Станиславовича, на диссертационную работу Тарасова Никиты Андреевича на тему: «Гибридные нейросетевые методы анализа понятности текстов юридических документов на русском языке» представленную к защите на соискание ученой степени кандидата технических наук по 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**

Использование автоматических методов обработки текстовых данных позволяют решать широкий спектр задач анализа юридических документов. Использование современных нейросетевых методов позволяет эффективно решать задачи, связанные с анализом больших корпусов правовых текстов. Интеллектуальные методы анализа текстовых данных дают возможность структурировать, систематизировать и категоризировать документы по различным языковым признакам. В частности, решение задачи определения сложности и доступности восприятия правовых текстов позволяет предотвращать двоечтения, неверное толкование официальных документов и установить эффективный язык коммуникации между юристами и лицами, не являющихся профессионалами в юридической сфере. Следует отметить, что переход к цифровым и интеллектуальным технологиям является приоритетом научно-технологического развития России.

### **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**

Научная новизна представленной работы заключается в разработке гибридного подхода оценки понятности юридических документов на русском языке, объединяющем классические методы ручного построения набора языковых характеристик с использованием современных языковых моделей. Были созданы и проанализированы базы данных юридических текстов различных типов и жанров. Ряд методов комплекса был дополнительно апробирован на реальных данных ответов на юридические запросы.

### **Теоретическая и практическая значимость работы**

Исследование и его результаты имеют теоретическую и практическую значимость. Разработан комплекс методологических и алгоритмических подходов и программных инструментов для оценки сложности и доступности восприятия юридических документов. На основе статистических методов, методов машинного обучения был выявлен набор характеристик, наиболее эффективно моделирующие сложность юридических текстов. В совокупности с современными нейросетевыми языковыми моделями был разработан программный комплекс для оценки сложности правовых текстов. Разработанные методы помогут повысить эффективность принятия решений профильных специалистов, а также находят применение в юридической и лингвистической научных сферах. В масштабах государства созданный подход способен как улучшить уровень взаимодействия с населением в законодательных процессах так и повысить уровень юридической грамотности.

### **Анализ содержания работы**

Во введении представлены основные сведения о диссертационной работе: 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 рассматривается способ оценки частотности использования юридических терминов в различных типах документов с использованием статистических методов. Приводится обзор существующих подходов к анализу частотных характеристик, производится сравнение характеристик различных широко используемых текстовых корпусов.

Во второй главе представлен набор признаков, которые определяют юридические документы с точки зрения их понятности. Проводится анализ эффективности этих признаков, и предлагается методика их применения для классификации документов по степени сложности. Проводится сравнение моделей, основанных на этих характеристиках, с алгоритмами, построенными на основе языковых моделей.

В третьей главе исследования представлен гибридный метод оценки сложности, который объединяет языковые характеристики и большие языковые модели. Использование языковых как части программного решения позволяет эффективно анализировать тексты на естественном языке, в особенности тексты с нестандартной структурой и языком.

В четвертой главе проводится сравнительный анализ сложности юридических документов различных типов и стилей с применением гибридной семантической модели. Приводится методология сбора, предварительной обработки и анализа документов, полученных из различных источников.

В пятой главе приводится пример практического использования части полученных методов. Анализ проводится на текстах вопросов и ответов на юридические запросы в налоговой сфере. Показано, как модели и методы возможно адаптировать для решения смежных задач.

### 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.

Кандидат технических наук,

Доцент с возложенными обязанностями зав. каф. технологии программирования СПбГУ,

19804, Санкт-Петербург, Университетский пр., 35,

Тел. (812) 428-71-59; e-mail: i.blekanov@spbu.ru

Блеканов Иван Станиславович



13.08.2024

