

## **ОТЗЫВ**

члена диссертационного совета доктора технических наук Баденко Владимира Львовича 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**

Применение автоматизированных методов обработки текстов на русском языке позволяет эффективно решать задачи анализа юридических документов, в частности за счет использования современных нейросетевых подходов применительно к обработке больших объемов текстов, в том числе правовых. Такие методы, в частности, позволяют структурировать и систематизировать правовые документы, опираясь на их языковые особенности. Определение сложности и доступности восприятия правовых текстов играет ключевую роль в предотвращении неверных интерпретаций и установлении эффективной коммуникации в юридической среде. Такие исследования находятся в русле цифровой трансформации экономики и развития интеллектуальных технологий, которые являются приоритетными направлениями в научно-технологическом развитии 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 работы заключается в создании комплексной гибридной нейросетевой методологии для задач оценки сложности и доступности восприятия юридических документов, которая включает методы семантического анализа правовых документов, с использованием языковых моделей и классического анализа языковых характеристик. Данная гибридная методология была протестирована на большом количестве разнообразных юридических документов.

33-06-1210 от 10.12.2024



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

Данное исследование и его результаты являются значимыми как в теоретическом, так и в практическом смысле. Разработаны методологии и алгоритмы, а также программные средства для анализа сложности восприятия правовых документов. С применением статистических методов и машинного обучения был выявлен набор характеристик, моделирующих сложность юридических текстов. С использованием современных нейросетевых языковых моделей создан инструмент для оценки сложности правовых текстов. Данные разработки способствуют улучшению процесса принятия решений профессионалами, а также могут быть полезны в дальнейших юридических и лингвистических исследованиях.

### **Анализ содержания работы**

Введение содержит следующие 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 исследуется метод оценки частотности юридических терминов. Во второй главе исследуются признаки, определяющие сложность юридических текстов, и разрабатывается методика классификации их сложности. Третья глава вводит гибридный метод оценки сложности, объединяющий языковые характеристики и большие языковые модели. В четвертой главе сравнивается сложность юридических текстов разных стилей с использованием гибридной модели, описываются способы сбора и анализа правовых документов. В пятой главе представлен пример практического применения разработанных методов для текстов ответов на запросы в налоговой сфере, демонстрируется их адаптируемость для решения сопутствующих задач.

### **Замечания**

По тексту диссертации имеются следующие замечания:

- Недостаточно много внимания уделяется описанию метода и модели, представленной в Главе 3 «Гибридная модель оценки сложности: разработка и применение для российских юридических текстов».
- Автору следовало бы четко сформулировать из чего состоит «Комплекс современных гибридных нейросетевых методов...», который является частью положений, выносимых на защиту.
- Отсутствуют формулы для используемых индексов сложности: Флеша-Кинкейда, SMOG, ARI и т.д.
- В чем именно состоят методологические основы адаптации больших языковых моделей в задаче определения понятности юридических текстов?
- Часть результирующих метрик работы, как в таблицах, так и на картинках представлены на английском языке, что затрудняет восприятие результатов.
- Отсутствует полное описание алгоритмов тестирования качества и процесса применения алгоритмов для анализа новых данных, а также блок-схема комплекса программ для информационного обеспечения научно-исследовательской и опытной работы с русскоязычными текстовыми данными правовых процессов.



### 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 Высшей школы гидротехнического и энергетического строительства СПбПУ, 195251, Санкт-Петербург, Политехническая ул., 29 Гидрокорпус-1, СПбПУ,

E-mail: badenko\_vl@spbstu.ru

Баденко Владимир Львович

Дата 26.11.2024

Подпись

