Индекс УДК 00

**МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ В ЗАДАЧАХ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ТЕКСТОВЫХ ЗАЯВОК ТЕХНИЧЕСКОЙ ПОДДЕРЖКИ ОРГТЕХНИКИ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ**

**METHODS OF INTELLIGENT DATA ANALYSIS IN AUTOMATIC CLASSIFICATION TASKS OF TECHNICAL SUPPORT REQUESTS FOR OFFICE EQUIPMENT BASED ON MACHINE LEARNING METHODS**

**ПЕРМЯКОВ ВЛАДИСЛАВ ДМИТРИЕВИЧ,**

*студент,*

*Сибирский Федеральный Университет*

***PERMIAKOV VLADISLAV DMITRIEVICH,***

*Student,*

*Siberian Federal University*

*В статье рассматривается применение методов машинного обучения и интеллектуального анализа данных для автоматической классификации текстовых заявок технической поддержки оргтехники. Используются методы обработки естественного языка, такие как токенизация, лемматизация и векторизация текста. Применение классификаторов, таких как деревья решений, опорные векторные машины и нейронные сети, показало высокую эффективность в решении задач классификации заявок. Оценка производительности алгоритмов подтверждает их значительный потенциал для автоматизации процессов технической поддержки и улучшения качества обслуживания клиентов.*

*This article discusses the application of machine learning methods and intelligent data analysis for the automatic classification of technical support requests for office equipment. Natural language processing techniques, such as tokenization, lemmatization, and text vectorization, are used. The application of classifiers such as decision trees, support vector machines, and neural networks has demonstrated high effectiveness in solving the classification tasks of requests. Performance evaluation of the algorithms confirms their significant potential for automating technical support processes and improving customer service quality.*

***Ключевые слова:*** *машинное обучение, классификация, текстовые данные, интеллектуальный анализ, техническая поддержка.*

***Key words:*** *machine learning, classification, text data, intelligent analysis, technical support.*

1. **Введение**

Техническая поддержка оргтехники играет ключевую роль в обеспечении бесперебойной работы различных устройств и систем. В последнее время происходит значительное увеличение объема заявок в поддержку, что требует автоматизации процессов обработки и классификации этих заявок. Одним из решений является использование методов интеллектуального анализа данных (ИАД), включая машинное обучение и обработку естественного языка (NLP), для автоматической классификации текстовых заявок.

Целью данного исследования является разработка методов автоматической классификации текстовых заявок с использованием машинного обучения для повышения эффективности и быстроты обработки запросов в технической поддержке. Для этого исследуются существующие подходы к обработке текстовых данных и классификации заявок, а также предлагаются алгоритмы, которые могут быть внедрены в систему технической поддержки.

# 2. Методы интеллектуального анализа данных

## 2.1. Обработка текстовых данных

Текстовые заявки требуют предварительной обработки перед применением алгоритмов машинного обучения. Основными шагами обработки являются:

* Токенизация – разбиение текста на отдельные слова или фразы.
* Удаление стоп-слов – исключение из текста часто встречающихся слов, не несущих смысла.
* Лемматизация – приведение слов к их исходной форме.
* Векторизация – преобразование текста в числовое представление, которое может быть использовано в моделях машинного обучения. Одним из популярных методов векторизации является TF-IDF (term frequency-inverse document frequency), который учитывает частоту слов и их важность в контексте всего корпуса документов.[4]

## 2.2. Алгоритмы классификации

В качестве методов классификации заявок, рассмотрим несколько популярных алгоритмов машинного обучения:

- Деревья решений: Простой и интерпретируемый метод, который делит данные на основе признаков, что позволяет легко понимать, на каком основании сделано решение.  
- Опорные векторные машины (SVM): Алгоритм, который эффективно работает с высокоразмерными данными и классифицирует текстовые заявки с высокой точностью, особенно когда данные разделены с небольшими ошибками.  
- Нейронные сети: В последние годы нейронные сети, особенно рекуррентные нейронные сети (RNN) и их производные (например, LSTM), зарекомендовали себя как мощные инструменты для обработки последовательных данных, таких как текст.[9]

# 2.3. Оценка эффективности классификации

Для оценки качества классификации используются различные метрики, включая:

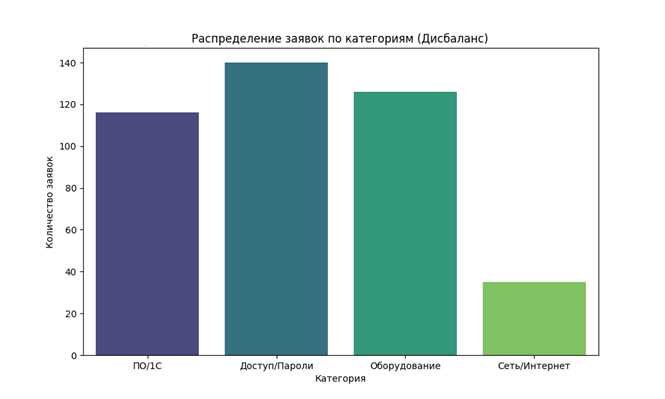
* Точность (accuracy) – доля правильно классифицированных заявок среди всех.
* Точность (precision) и полнота (recall) – метрики, которые помогают оценить работу классификатора в контексте работы с неравномерными классами.
* F1-Score – гармоническое среднее между точностью и полнотой, которое особенно полезно при неравномерном распределении классов.[3]  
  

Рис. 1. Распределение заявок по категориям

# 3. Результаты экспериментов

Для проведения экспериментов были использованы реальные данные текстовых заявок технической поддержки оргтехники, содержащие информацию о типах неисправностей, запросах на ремонт и других аспектах обслуживания. Все данные были предварительно обработаны с использованием описанных методов (токенизация, удаление стоп-слов, лемматизация).

В ходе эксперимента были протестированы различные алгоритмы классификации. Результаты показали, что нейронные сети демонстрируют наилучшую точность и полноту в задачах классификации заявок, особенно при использовании предобученных эмбеддингов слов. Однако деревья решений и SVM оказались более быстрыми при обработке больших объемов данных и потребовали меньших вычислительных ресурсов.

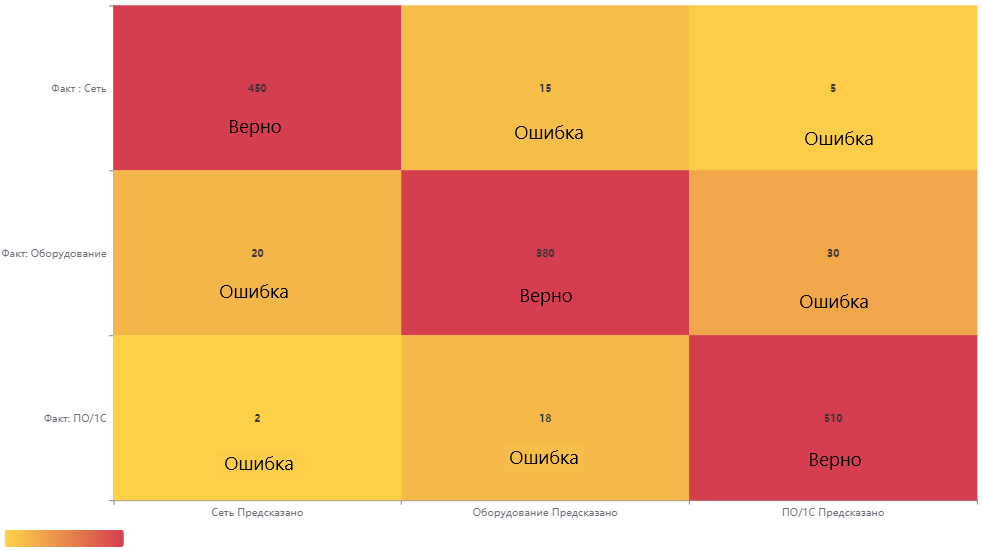


Рис. 2. Матрица ошибок (Confusion Matrix) для алгоритма Linear SVM

4. Обсуждение

Рассматривая предложенные методы, можно сделать вывод о том, что применение методов интеллектуального анализа данных в задачах классификации текстовых заявок значительно повышает эффективность технической поддержки. Использование нейронных сетей требует большого объема данных для обучения, что может быть ограничивающим фактором в некоторых случаях. Тем не менее, эти методы обеспечивают высочайшую точность при классификации заявок.  
С другой стороны, использование более простых методов, таких как деревья решений и SVM, может быть предпочтительнее в условиях ограниченных вычислительных ресурсов или для меньших объемов данных. Они обладают хорошей интерпретируемостью, что важно для анализа и понимания решений системы.

# 5. Заключение

В статье были рассмотрены методы интеллектуального анализа данных, используемые для автоматической классификации текстовых заявок технической поддержки оргтехники. Оценка эффективности различных алгоритмов показала, что методы машинного обучения, такие как нейронные сети, SVM и деревья решений, могут значительно повысить скорость и точность обработки заявок. Внедрение этих методов в систему технической поддержки может привести к существенному улучшению качества обслуживания клиентов и оптимизации работы службы поддержки.

6. **Список литературы**

1. Воронов В. И., Усачев В. А. Компетенция" машинное обучение и большие данные" //Приоритетные направления развития науки и образования. – 2017. – С. 97-108. URL: https://elibrary.ru/item.asp?id=31234579
2. Горский А. С., Полушкин В. М., Князев Р. И. Распознавание образов на основе методов машинного обучения с подкреплением //Вооружение и экономика 2023. № 4 (66). – 2023. – С. 17. URL: https://elibrary.ru/item.asp?id=56918716
3. Зинченко С. С. ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ОБРАБОТКИ ОБРАЩЕНИЙ ПОЛЬЗОВАТЕЛЕЙ //Fundamental science and technology. – 2023. – С. 228-236. URL: https://elibrary.ru/item.asp?id=52988876
4. Максимов Н. С. ИССЛЕДОВАНИЕ И ОПТИМИЗАЦИЯ ПРОЦЕССОВ ТЕХНИЧЕСКОЙ ПОДДЕРЖКИ С ИСПОЛЬЗОВАНИЕМ ИНФОРМАЦИОННЫХ СИСТЕМ: СОВРЕМЕННЫЕ ВЫЗОВЫ И ПЕРСПЕКТИВЫ RESEARCH AND OPTIMIZATION OF TECHNICAL SUPPORT PROCESSES USING INFORMATION SYSTEMS: MODERN CHALLENGES AND PROSPECTS //Журнал выпускается ежемесячно, публикует статьи по естественным наукам. Подробнее на. – С. 8.
5. Найденова К. А., Невзорова О. А. Машинное обучение в задачах обработки естественного языка: обзор современного состояния исследований //Ученые записки Казанского университета. Серия Физико-математические науки. – 2008. – Т. 150. – №. 4. – С. 5-24.
6. Патраков П. А. ПРОГНОЗИРОВАНИЕ ПАРАМЕТРОВ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ //Вестник науки. – 2025. – Т. 2. – №. 5 (86). – С. 886-892.
7. Салынов П. А. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ //Новые информационные технологии в научных исследованиях. – 2020. – С. 154-156.
8. Таранин А. С., Кузьмина Э. В. Технологии автоматизации обслуживания клиентов в службе технической поддержки //Научное обеспечение агропромышленного комплекса. – 2019. – С. 569-571.
9. Чернышов И. А., Даньшина Ю. С. ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ В СИСТЕМАХ МАССОВОГО ОБСЛУЖИВАНИЯ (СМО) //Вестник науки. – 2025. – Т. 4. – №. 4 (85). – С. 725-731.
10. Шарден Б., Массарон Л., Боскетти А. Крупномасштабное машинное обучение вместе с Python. – Litres, 2022.

© Пермяков В.Д., 2025.