

## ОБЗОР АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБРАБОТКИ ЗАПРОСОВ В HELPDESK

**Аннотация:** статья представляет собой комплексное исследование современных методов применения машинного обучения в области обслуживания клиентов и поддержки пользователей через helpdesk. Рассматриваются разнообразные алгоритмы машинного обучения, такие как методы классификации и кластеризации, применяемые для эффективной обработки запросов и улучшения процессов поддержки.

**Ключевые слова:** автоматизация, машинное обучение, служба технической поддержки, обработка входящих запросов, методы классификации.

**Annotation:** the article represents a comprehensive study of contemporary methods for applying machine learning in the field of customer service and user support through helpdesk. Various machine learning algorithms, such as classification and clustering methods, are examined for efficient handling of requests and improvement of support processes.

**Keywords:** automation, machine learning, technical support service, processing of incoming requests, classification methods.

В условиях постоянного роста клиентской базы и объема данных в сфере информационных технологий, эффективная обработка запросов в helpdesk подразделениях становится жизненно важной для обеспечения качественного

обслуживания клиентов. Методом оптимизации этого бизнес-процесса может стать автоматизация с использованием машинного обучения.

В целях улучшения эффективности процесса обработки и распределения входящих запросов в helpdesk необходимо провести анализ существующих моделей обработки данных и выявить наилучший метод. Этот анализ может включать в себя оценку параметров, таких как скорость обработки, использование ресурсов и точность результатов. После выделения оптимальной модели требуется реализовать программное обеспечение, основанное на данной модели, и провести экспериментальное подтверждение повышения эффективности в реальных условиях функционирования системы обработки запросов. Полученные результаты будут служить основой для принятия обоснованных решений по улучшению общей производительности системы.

В работах авторов таких, как Батура Т.В. («Методы автоматической классификации текстов») [1], Комарова А.В. («Метод автоматизированного извлечения адресов из неструктурированных текстов») [2], Кашницкий Ю.С. («Ансамблевый метод машинного обучения, основанный на рекомендации классификаторов») [3], Кузнецов И.А. («Методы и алгоритмы машинного обучения для предобработки и классификации слабоструктурированных текстовых данных в научных рекомендательных системах») [4], осуществляется анализ разнообразных подходов к автоматической обработке и классификации текстов. Внимание уделяется детальному изучению особенностей алгоритмов машинного обучения, применяемых для классификации текстовых документов и данных. Особый фокус направлен на рассмотрение ансамблевых методов машинного обучения, представляющих собой стратегию объединения нескольких базовых моделей с целью достижения более высокой точности и стабильности в процессе классификации текстов. Однако, в данных исследованиях не проводился

подробный анализ применения упомянутых методов и алгоритмов в контексте обработки входящих запросов в подразделениях helpdesk.

В области машинного обучения традиционно выделяются три широкие категории методологических подходов:



***Рисунок 1. Методы машинного обучения***

1. **Обучение под присмотром (с учителем):** В данной парадигме предоставляются компьютеру примеры входных данных, снабженные соответствующими желаемыми результатами, поставляемыми "учителем". Задача заключается в формировании общего правила, которое абстрагирует связь между предоставленными входными данными и соответствующими выходными результатами.
2. **Обучение без учителя:** В этом контексте алгоритму обучения не предоставляются метки, и ему требуется самостоятельно выявлять структуру во входных данных. Обучение без учителя может служить самоцелью, когда целью является выявление скрытых закономерностей в данных, либо использоваться в качестве средства достижения определенных целей, таких как обучение функций.
3. **Обучение с подкреплением:** В этом подходе компьютерная программа взаимодействует с динамическим окружением, в котором ей предстоит

достичь конкретной цели, например, такой как управление транспортным средством или игра против оппонента. По мере передвижения по пространству задачи программа получает обратную связь, аналогичную вознаграждению, стремясь максимизировать получаемую положительную реакцию.

Рассмотрим в таблице 1 плюсы и минусы этих подходов:

*Таблица 1*

<b>Подход</b>	<b>Плюсы</b>	<b>Минусы</b>
Обучение под присмотром (с учителем)	<p><b>Точность</b> (при наличии размеченных данных этот подход обеспечивает высокую точность классификации запросов)</p> <p><b>Стабильность</b> (хорошо работает в случаях, когда имеется набор четко определенных классов запросов)</p> <p><b>Интерпретируемость</b> (легче интерпретировать и понять результаты модели, так как она обучается на основе предоставленных примеров)</p>	<p><b>Требование к размеченным данным</b> (требует большого объема размеченных данных для эффективного обучения)</p> <p><b>Зависимость от качества данных</b> (результаты сильно зависят от качества и представительности обучающих данных)</p>
Обучение без учителя	<p><b>Способность обнаруживать скрытые закономерности</b> (подходит для обработки запросов с неизвестной структурой, позволяет выявлять скрытые закономерности)</p> <p><b>Без необходимости в размеченных данных</b> (может обрабатывать данные без предварительной разметки, что особенно полезно в случаях, когда доступ к размеченным данным ограничен)</p>	<p><b>Неопределенность в результатах</b> (модель может выдавать менее точные результаты из-за отсутствия конкретных меток классов)</p> <p><b>Трудность интерпретации</b> (сложнее интерпретировать внутреннюю структуру модели и ее выводы)</p>
Обучение с подкреплением	<p><b>Адаптивность к изменениям</b> (модель обучается на основе обратной связи, что позволяет ей адаптироваться к изменяющимся требованиям и характеру запросов)</p> <p><b>Применимость к динамическим сценариям</b> (эффективен в обработке запросов, изменяющихся во времени, например, при обслуживании клиентов)</p>	<p><b>Трудность определения подходящей целевой функции</b> (выбор целевой функции может быть нетривиальным, что затрудняет оптимизацию модели)</p> <p><b>Высокие требования к вычислительным ресурсам</b> (обучение с подкреплением может требовать значительных вычислительных мощностей)</p>

Хотя каждый алгоритм имеет свои преимущества и ограничения, ни один алгоритм не является универсальным, подходящим для любой задачи.

Наиболее подходящими алгоритмы для решения задач исследования, выявлены следующие:

1. Классификация может быть эффективным методом, если задача состоит в определении типа заявки (например, вопросы, проблемы, запросы на информацию).
2. Кластеризация, если необходимо группировать похожие заявки для их эффективной обработки. Кластеризация может быть полезна для выявления общих тем или проблем.
3. Обучение с учителем в условиях несбалансированных данных в случае, когда классы заявок представлены неравномерно, и необходимо учесть дисбаланс для более точного обучения модели.
4. Обучение с подкреплением, при возможности сформулировать процесс принятия решений в службе поддержки в виде задачи обучения с подкреплением. Например, если важно обучить систему принимать определенные действия на основе опыта.
5. Нейронные сети и глубокое обучение, когда имеется доступ к большому объему данных и задача требует обработки сложных структур или текстовой информации.
6. Метод ближайших соседей (KNN) и метод опорных векторов (SVM). В случае небольших объемов данных, где эти методы могут работать достаточно хорошо, и нет необходимости в использовании сложных моделей.
7. Ансамблевые методы: Ансамблевые методы, такие как случайные леса или градиентный бустинг, могут быть эффективными, когда нужно комбинировать предсказания нескольких моделей для улучшения общей производительности.

8. Формирование признаков. Если важно автоматически извлекать и выделять важные характеристики из данных, особенно при работе с неструктурированным текстом.

Выбор метода машинного обучения для обработки входящих заявок в службу поддержки зависит от множества факторов, включая характер данных, объем информации, доступность размеченных данных, требования к интерпретируемости модели и степень сложности задачи. В некоторых случаях может быть полезно использовать несколько методов в комбинации для достижения лучших результатов.

#### **Использованные источники:**

1. Батура Т.В. Методы автоматической классификации текстов // Программные продукты и системы. – Т. 30, № 1, 2017.
2. Комарова А.В., Менщиков А.А., Полев А.В., Гатчин Ю.А. Метод автоматизированного извлечения адресов из неструктурированных текстов // International Journal of Open Information Technologies. – Т.5, № 11, 2017.
3. Кашницкий Ю.С., Игнатов Д.И. Ансамблевый метод машинного обучения, основанный на рекомендации классификаторов // Интеллектуальные системы. Теория и приложения. – Т.19, № 4, 2015.
4. Кузнецов И.А. Методы и алгоритмы машинного обучения для предобработки и классификации слабоструктурированных текстовых данных в научных рекомендательных системах // АВТОРЕФЕРАТ диссертации на соискание ученой степени к.т.н. – Москва – 2019.