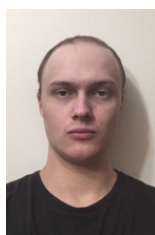


УДК 004.042

## DATA ANALYTICS В ЧАТАХ



**Ю.П. Плахотникова**  
студент БГУИР  
pocketickchaya@gmail.com



**Т.С. Ширяев**  
студент БГУИР  
tim.shiryaev@gmail.com



**С.Н. Нестеренков**  
доцент  
кафедры программного обеспечения  
информационных технологий, декан  
факультета компьютерных систем  
и сетей, кандидат технических наук  
s.nesterenkov@bsuir.by

### **Ю.П. Плахотникова**

Студент 4 курса специальности “Информатика и технологии программирования” БГУИР.

### **Т.С. Ширяев**

Студент 4 курса специальности “Информатика и технологии программирования” БГУИР.

### **С.Н. Нестеренков**

Кандидат технических наук, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры Программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

**Аннотация.** Рассмотрены методы анализа данных в чатах, которые используются для изучения и оценки продукта в целях его совершенствования. Описывается процесс сбора данных, их обработки и представления. Описаны принципы и технологии машинного обучения, которые можно применять для анализа данных в чатах и их преимущества. С помощью примеров из реальной жизни показывается, как анализ данных в чатах может быть полезен для бизнеса.

**Ключевые слова:** анализ данных, машинное обучение, чаты.

### **Введение.**

Анализ данных – это процесс изучения, очистки, преобразования и моделирования данных с целью получения полезной информации и поддержки принятия решений. Этот процесс включает в себя использование математических методов, статистических алгоритмов и компьютерных технологий. Анализ данных — это ключевой элемент многих областей деятельности, например, в бизнесе, медицине, науке и общественных науках.

В бизнесе анализ данных часто используется для прогнозирования спроса на товары и услуги, определения цен и маркетинговых стратегий. Для анализа данных в бизнесе может использоваться множество инструментов, включая макросы Excel, программное обеспечение для бизнес-аналитики, базы данных и инструменты машинного обучения.

Современными способами общения являются чаты в мессенджерах, социальных сетях и других приложениях. Анализ данных в чатах играет огромную роль в современных коммуникациях. Это связано с тем, что каждый день миллионы людей по всему миру используют чаты для онлайн-общения, работы, обучения и даже для прохождения медицинских консультаций. Многие из этих сообщений содержат важную информацию, которая может быть использована для анализа, определения различных показателей и получения ценной информации. Анализ данных в чатах — это процесс извлечения информации из текстовых сообщений, которые отправляют пользователи. Это может

включать в себя анализ тональности, определение ключевых слов и тематики, анализ эмоций и т.д. Анализ тональности позволяет определить, каким образом пользователи выражают свои эмоции и отношения к определенным событиям. Например, если пользователи обсуждают новинку, то можно узнать, каковы их отзывы - положительные, отрицательные или нейтральные. Анализ тональности может дать представление общего отношения к бренду, товару или услуге, что может использоваться для принятия решений по улучшению продукта или коммуникации с клиентами.

Ключевые слова и тематика также важны при анализе данных в чатах. Количество и частота использования определенных слов может дать представление об интересах и потребностях пользователей. Например, если пользователи часто упоминают слова "путешествия", "отпуск" и "туризм", то можно сделать вывод, что у них есть интерес к этим темам. Данные могут быть использованы для персонализации сообщений и предложений, чтобы лучше соответствовать потребностям пользователей. Анализ эмоций позволяет выявить какие эмоции вызывает определенный продукт, событие или бренд. Эта информация может быть использована, чтобы сделать более эффективные маркетинговые кампании и улучшить общий имидж компании.

Важно понимать, что анализ данных в чатах может быть сложным и требует использования специальных инструментов и технологий. Однако, если правильно провести анализ и извлечь ценную информацию, то это может помочь улучшить качество продукта, снизить затраты и усилить взаимодействия с клиентами.

В целом, анализ данных в чатах — это мощный способ получения ценной информации о потребностях и интересах пользователей. Это может помочь компаниям улучшить свою продукцию и услуги, а также оптимизировать маркетинговые кампании.[1]

### **Оперативный анализ.**

По мере того, как компании начинают осознавать ценность анализа текстового контента, содержащегося в чате, они делают это двумя разными способами:

1. Анализ пользовательского чата. Эта методология анализа фокусируется на сообщениях, отправленных клиентом, чтобы раскрыть темы, возникающие проблемы, обсуждаемые клиентской базой.

2. Анализ чата пользователя с технической поддержкой. Этот метод анализа фокусируется на мониторинге поведения агентов поддержки, чтобы обеспечить соблюдение протоколов и определить области для улучшения ПО.

Разделив анализ на эти два разных способа, компании получают больше возможностей для улучшения деловой практики как внутри компании, так и за ее пределами. Также с точки зрения реализации анализ данных в чатах можно разделить на несколько категорий:

1. Использование алгоритмов машинного обучения. Алгоритмы машинного обучения способны проанализировать большой объем данных с высокой точностью. Они могут идентифицировать изображения, тексты и другой контент, чтобы получить полезную информацию. Модели машинного обучения умеют распознавать формат и структуру сообщений, а также делать выводы на основе анализа текста.

2. Анализ данных в режиме реального времени. Анализ данных в режиме реального времени позволяет мгновенно получать информацию о чатах по мере их появления. Режим реального времени актуален, когда нужно реагировать на данные мгновенно. Этот способ анализа данных может повысить эффективность работы чатов. Системы, работающие в режиме реального времени, могут быстро и эффективно обрабатывать даже огромные объемы данных и выдавать результаты анализа в мгновение ока. Исключительно важно, чтобы данные поступали в систему быстро, иначе система не сможет реагировать корректно.

3. Анализ, включающий в себя работу эмоционального интеллекта. Эмоциональный интеллект — это способность человека понимать свои и чужие эмоции, управлять ими и использовать в процессе общения. Аналитика на основе эмоционального интеллекта может быть использована для идентификации сильных и слабых сторон в коммуникации пользователей в чате. Она может обратить внимание на нарушения, убыль эмоциональной интенсивности, перебои в общении и необходимость корректировки манеры общения, что может помочь в разработке ИИ.

#### **Центральные темы.**

Лучшее место для начала анализа текста данных чата — это определение основных тем с помощью комбинаций ключевых слов и категорий. Ключевые слова — это парные сочетания слов, определяемые по частоте их употребления. Форматы облака тегов обеспечивают простую визуализацию данных для определения наиболее часто упоминаемых терминов в контенте. С помощью облаков тегов легко определить, какие типы продуктов или названия (например, Money Market или Debit Card) часто обсуждаются. Категории (типы) разговоров автоматически определяются платформой на основе множества факторов, включая частоту, силу и качество словесных ассоциаций в данных чата, настроение и другие. Они отображаются в порядке статистической значимости. Визуализация как категорий, так и комбинаций ключевых слов позволяет отобразить наиболее важные области, требующие внимания.

#### **Тенденции.**

Следующим шагом в анализе данных чата является поиск тенденций в каждой категории и в наиболее часто используемых словосочетаниях. Важно знать, повышается или понижается объем и общее впечатление по каждой теме месяц за месяцем. Например, финансовый отдел может столкнуться с ошибками, возникающими при использовании продукта, и ростом ряда вопросов, связанных с налогообложением, весной, достигающими пика в апреле. Тенденции текстовой аналитики также могут определять вопросы, возникающие при использовании продукта, и комментарии о его обновлениях. Возможно, была запущена новая версия мобильного приложения, что вызвало увеличение количества обсуждений программного обеспечения. Предприятия могут внедрять планы совершенствования процессов и ожидать от клиентов подтверждения того, что их усилия принесут желаемые результаты. В любом случае тенденции данных в диалогах покажут прогресс в разработке продукта.

#### **Постоянные проблемы.**

Текстовая аналитика чата также может выявить проблемы, которые регулярно возникают у клиентов. Примеры включают часто задаваемые вопросы, которые не рассматриваются на странице часто задаваемых вопросов веб-сайта. Могут быть трудности с определенными версиями браузера или операционной системы, вызывающие разное взаимодействие с пользователем. Стандартный процесс бизнес-транзакций может быть источником поддержки в чате. Изучая данные сеансов чата, можно выявить эти постоянные проблемы и глубже изучить переменные, такие как географическое распределение. Изучение отдельных сеансов чата поможет подтвердить догадки и достичь лучших результатов. Также важно искать «социальных лидеров мнений» в данных диалогов и убедиться, что их проблемы решены на 100% [2].

В качестве улучшения программного обеспечения на основе полученных в ходе анализа данных можно обеспечить:

- Дополнительные функции/возможности в продукте
- Обновление часто задаваемых вопросов на сайте
- Улучшение пользовательского интерфейса
- Улучшение обмена сообщениями об ошибках и обработки ошибок
- Корректировка взаимодействия сотрудников с клиентской базой

— Внедрение перспективного плана анализа данных для выявления проблем практически в реальном времени.

**Заключение.**

Анализ набора неструктурированных данных сеансов чата может напрямую привести к повышению эффективности бизнеса. Текстовая аналитика не только выявит конкретные проблемные области, но и подтвердит, а также обеспечит успешные обновления ПО.

**Список литературы**

[1] Вайгенд Андреас. BIG DATA. Вся технология в одной книге / Андреас Вайгенд. – Москва: Эксмо, 2018.–384 с.

[2] Радченко, И.А. Технологии и инфраструктура Big Data / И.А. Радченко, И.Н. Николаев. – СПб: Университет ИТМО, 2018.–52 с.

**DATA ANALYTICS IN CHATS**

***Y.P. Plakhotnikova***  
*Student of BSUIR*

***T.S. Shiryayev***  
*Student of BSUIR*

***S.N. Nesterenkov***  
*Candidate of Technical Sciences,*  
*Associate Professor of the Department*  
*of Information Technology Software,*  
*Dean of the Faculty of Computer*  
*Systems and Networks*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus*  
*E-mail: packetickchaya@gmail.com, tim.shiryayev@gmail.com, s.nesterenkov@bsuir.by*

**Abstract.** The methods of data analysis in chats, which are used to study and evaluate the product in order to improve it, are considered. The process of data collection, processing and presentation is described. The principles and technologies of machine learning that can be used to analyze data in chats and their advantages are described. Using real-life examples, it shows how data analysis in chats can be useful for business.

**Keywords:** data analysis, machine learning, chats.