



Анализ отзывов пациентов с использованием машинного обучения и лингвистических методов

© 2025, И.Е. Калабихина¹, В.С. Мошкин²✉, А.В. Колотуша¹, М.И. Кашин², Г.А. Клименко¹, З.Г. Казбекова¹

¹Московский государственный университет имени М.В. Ломоносова (МГУ), Москва, Россия

²Ульяновский государственный технический университет (УлГТУ), Ульяновск, Россия

Аннотация

С развитием цифровизации традиционные методы анкетирования потребителей с целью оценки степени их удовлетворённости качеством услуг уступают место подходу, основанному на автоматической обработке текстовых массивов социальных медиа. Целью работы является определение степени удовлетворённости качеством медицинских услуг пациентов посредством разработки и апробации алгоритма классификации русскоязычных текстовых отзывов, извлечённых из социальных медиаресурсов. Интерес представляет определение тональности отзывов пациентов (положительный/отрицательный) о работе медицинских учреждений и врачей, а также объекты обращения отзыва – качество оказанных медицинских услуг или организация обслуживания пациентов медицинским учреждением. Разработан метод классификации текстовых отзывов о работе медицинских учреждений, размещённых пациентами на двух сайтах отзывов о врачах в России. Проанализировано около 60 тысяч отзывов. Апробированы методы машинного обучения с использованием различных архитектур искусственных нейронных сетей. Разработанный алгоритм классификации имеет высокую эффективность – лучший результат показала архитектура на основе рекуррентной нейронной сети (показатель точности = 0.9271). Применение метода поиска именованных сущностей к текстовым сообщениям позволило повысить эффективность классификации для каждого из классификаторов, базирующихся на использовании нейронных сетей. Для повышения качества классификации требуется семантическое разбиение отзыва по объекту обращения и тональности и последующий учёт полученных фрагментов отдельно друг от друга.

Ключевые слова: машинное обучение, отзывы пациентов, нейронные сети, классификация отзывов, качество медицинских услуг.

Цитирование: Калабихина И.Е., Мошкин В.С., Колотуша А.В., Кашин М.И., Клименко Г.А., Казбекова З.Г. Анализ отзывов пациентов с использованием машинного обучения и лингвистических методов. *Онтология проектирования*. 2025. Т.15, №1(55). С.55-66. DOI:10.18287/2223-9537-2025-15-1-55-66.

Финансирование: Работа выполнена при финансовой поддержке Российского научного фонда, проект №23-71-01101 «Разработка моделей и методов повышения производительности хранилищ данных посредством предиктивного анализа темпоральной диагностической информации» и НИР «Воспроизводство населения в социально-экономическом развитии» при поддержке экономического факультета МГУ имени М.В. Ломоносова (ЦИТИС: 122041800047-9 ЭФ МГУ).

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Введение

В настоящее время процесс анкетирования для оценки удовлетворённости потребителей набором услуг начал уступать место подходу, основанному на автоматической обработке текстов в социальных медиа с возможностью извлечения семантики. При использовании такого подхода выборка становится более представительной и независимой, а результаты анализа более достоверно отражают отношение потребителей. Особенно важно проведение подобного анализа в таких сферах деятельности человека, как медицина.

Целью работы является определение степени удовлетворённости качеством медицинских услуг пациентов путём классификации русскоязычных текстовых отзывов, извлечённых из социальных медиаресурсов, о работе медицинских учреждений и врачей.

Для достижения данной цели необходима разработка программных классификаторов, позволяющих разделить анализируемые текстовые данные по группам в соответствии с выбранными критериями: эмоциональная окраска; объект или субъект применения; причинно-следственная составляющая и пр. Жанровые и речевые особенности текстов в социальных медиа порождают необходимость разработки интеллектуальных алгоритмов классификации текстов, позволяющих проводить подробный анализ текстовых отзывов с учётом особенностей предметной области.

Онлайн-отзывы и онлайн-рейтинги формируют *неформальные коммуникации, направленные на потребителей посредством интернет-технологий, связанные с использованием или характеристиками конкретных товаров и услуг или их продавцов* [1]. Это активно используется в здравоохранении для оценки деятельности врачей и больниц.

На отзыв пациента могут влиять личностные характеристики врача: пол, возраст, специальность [2–5]. Например, более высокие оценки у врачей-женщин [2, 3], акушеров-гинекологов и врачей более молодого возраста [6]. В зависимости от частоты использования веб-сайтов для рейтингования врачей [7, 8] пользователи, имеющие разные ключевых характеристики, по-разному оценивают значимость онлайн-отзывов о врачах [9]. В качестве данных используются одновременно значение рейтинга и тексты комментариев [10]. В частности, выявлены факторы, влияющие на положительные оценки врача, связанные с его характеристиками и другими, не зависящими от него факторами.

В ряде исследований используются в качестве основы данных массивы текстов отзывов о врачах [11, 12]. Оценки врачей в социальных медиа могут дополнять информацию, предоставляемую традиционными опросами пациентов, и способствовать пониманию пациентами качества услуг, предоставляемых врачом или медицинским учреждением [13].

Анализ социальных медиа включает:

- 1) извлечение из различных ресурсов соответствующего контента социальных сетей [14];
- 2) выбор данных для прогнозного моделирования анализа настроений;
- 3) визуализация выводов, полученных в результате анализа [15].

Для анализа настроений на основе полученных данных можно использовать контролируемые и неконтролируемые методы [16]. Основные подходы к классификации полярности анализируемых текстов включают слово, предложение или абзацы.

В [17] рассмотрены методы интеллектуального анализа текста, в т.ч. на основе машинного обучения и онтологий, а также с использованием гибридного подхода. Показано, что не существует алгоритма, который бы работал хорошо для всех типов данных.

В [18] рассмотрены различные типы классификаторов текста, в частности: байесовский классификатор; дерево решений; K -ближайшего соседа (K -NN); метод опорных векторов (SVM); искусственная нейронная сеть (НС) на основе многослойного перцептрона; алгоритм Роккио. Общим недостатком всех алгоритмов является их низкая производительность.

В [19] рассмотрены два подхода к классификации текста: подход «Мешок слов»; сетевая классификация. Выделено 18 классов и классифицированы актуальные темы. Показано, что сетевой классификатор работал значительно лучше, чем текстовый. В [20] обсуждаются методы, позволяющие преодолеть трудности классификации коротких текстов на основе потоковых данных в социальных сетях.

В [21] предложена многомерная структура классификации текстовых документов. Показано, что классификация на основе многомерной модели категорий с использованием многомерных и иерархических классификаций превосходит плоскую классификацию.

В статье [22] предложен метод точной настройки модели, обученной с использованием некоторых известных документов, содержащих более богатую контекстную информацию.

Как показал приведённый обзор, в настоящее время не существует единого подхода к классификации текстовых ресурсов. Результаты категоризации зависят от предметной области, репрезентативности обучающей выборки и других факторов. Поэтому актуальной является задача разработки и применения интеллектуальных методов анализа отзывов об оказании медицинских услуг.

1 Модели классификации отзывов о медицинских услугах

В настоящем исследовании разработан метод классификации текстовых отзывов, полученных из социальных медиа. Результатом классификации стало распределение текстовых отзывов по следующим основаниям:

- тональность текста: положительная или отрицательная;
- объект адресации: отзыв о медицинском учреждении или о конкретном враче.

Для классификации отзывов были апробированы методы машинного обучения с использованием различных архитектур НС. В данном исследовании предложено три архитектуры НС, которые зарекомендовали себя наилучшим образом в задачах небинарной классификации текстовых данных. Эффективность предложенных алгоритмов сопоставлена с результатами классификации текстов с использованием моделей, показывающих хорошие результаты при бинарной классификации (*BERT* и *SVM*) [23, 24].

1.1 Сеть долгой краткосрочной памяти

Общая архитектура *LSTM* (*Long Short-Term Memory*) – сети, представленная на рисунке 1, состоит из следующих слоёв:

- Слой *Embedding* – входной слой НС, состоящий из нейронов:
 $Emb = \{Size(D), Size(S_{vec}), L_{sec}\}$, где $Size(D)$ – размер словаря в текстовых данных;
 $Size(S_{vec})$ – размер векторного пространства, в которое будут вставлены слова; $Size(S_{vec}) = 32$;
 L_{sec} – длина входных последовательностей, равная максимальному размеру вектора, сформированного при предобработке слов.
- Слой *LSTM* – рекуррентный слой НС. Включает 32 блока.
- Слой *Dense* – выходной слой, состоящий из четырёх нейронов. Каждый нейрон отвечает за выходной класс. Функция активации – «*softmax*».

1.2 Рекуррентная НС

Общая архитектура рекуррентной НС, представленная на рисунке 2, состоит из следующих слоёв:

- Слой *Embedding* – входной слой НС.
- Слой *GRU* – рекуррентный слой НС. Включает 16 блоков.
- Слой *Dense* – выходной слой, состоящий из четырёх нейронов. Функция активации – «*softmax*».

1.3 Свёрточная НС

Общая архитектура свёрточной НС (*Convolutional neural network, CNN*), представленная на рисунке 3, состоит из следующих слоёв:

- Слой *Embedding* – входной слой НС.
- Слой *Conv1D* – свёрточный слой, необходим для глубокого обучения. С данным слоем точность классификации текстовых сообщений повышается. Функция активации – «*relu*».
- Слой *MaxPooling1D* – слой, отвечающий за уменьшение размерности сформированных карт признаков. Максимальный пул равен 2.
- Слой *Dense* – первый выходной слой, состоящий из 128 нейронов. Функция активации – «*relu*».
- Слой *Dense* – итоговый выходной слой, состоящий из четырёх нейронов. Функция активации – «*softmax*».

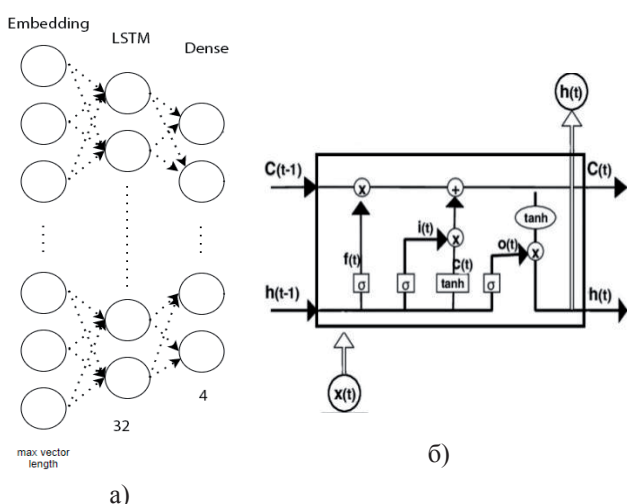


Рисунок 1 – LSTM-сеть: общая архитектура (а), LSTM-слой (б)

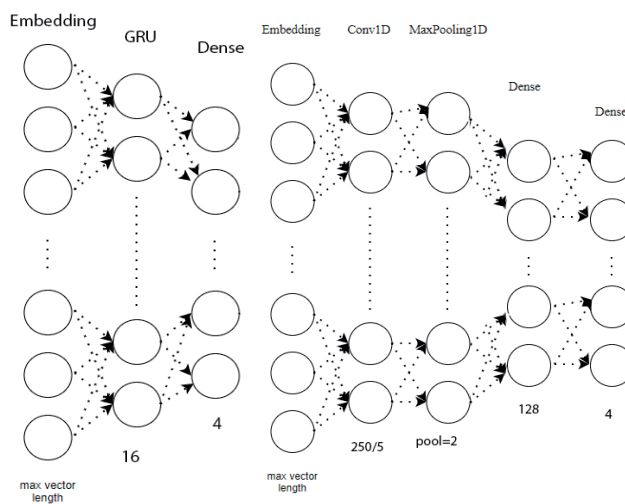


Рисунок 2 – Общая архитектура рекуррентной нейронной сети

Рисунок 3 – Общая архитектура сверточной нейронной сети

1.4 Использование лингвистических алгоритмов

Особенностью анализируемых текстовых отзывов является наличие элементов разных классов внутри одного отзыва. Для повышения качества классификации объединены наиболее эффективные методы машинного обучения с лингвистическими методами, которые учитывают речевые и грамматические особенности языка текста. Общая схема алгоритма классификации представлена на рисунке 4.

В качестве лингвистической составляющей разработанного алгоритма принята совокупность методов предобработки, валидации и детектирования именованных существ (имена врачей клиники).

2 Программная реализация системы классификации текстов

В качестве модуля для лингвистического анализа текста на русском языке использовалась библиотека *Natasha* на *Python*. С помощью данной библиотеки решаются базовые задачи обработки русского языка: сегментация на токены и предложения, морфологический и синтаксический анализ, лемматизация, извлечение, нормализация и детектирование именованных существ. Библиотека использовалась для поиска и извлечения именованных существ.

Для процессов инициализации, обучения НС, а также оценки эффективности классификации использовались следующие библиотеки:



Рисунок 4 – Общая схема алгоритма классификации

- *Tensorflow* – открытая программная библиотека для машинного обучения.
- *Keras* – библиотека глубокого обучения, представляющая высокоуровневый *API* на *Python*, способный работать поверх *TensorFlow*.
- *Numpy* – библиотека на *Python*, предназначенная для работы с многомерными массивами.
- *Pandas* – библиотека на *Python*, предоставляющая специальные структуры данных и операции для работы с числовыми таблицами и временными рядами.

Для обучения моделей использован *Google Colab*.

3 Результаты экспериментов по классификации текстовых отзывов

3.1 Набор данных

Для оценки эффективности предложенных подходов проведён ряд экспериментов по классификации текстовых отзывов о медицинских услугах клиник и врачей. В качестве исходных данных использовались текстовые отзывы агрегаторов *prodoctorov.ru* и *infodoctor.ru*.

Извлечённые данные имели следующий перечень переменных:

- *city* – город, в котором оставлен отзыв;
- *text* – текст отзыва;
- *author_name* – имя автора отзыва;
- *date* – дата написания отзыва;
- *day* – день написания отзыва;
- *month* – месяц написания отзыва;
- *year* – год написания отзыва;
- *doctor_or_clinic* – бинарная переменная (отзыв написан о враче ИЛИ о клинике);
- *spec* – специальность врача (для отзывов, посвящённых врачам);
- *gender* – пол автора отзыва;
- *id* – идентификационный номер отзыва.

По условиям экспериментов максимальное число слов в отзыве равно 90.

3.2 Результаты эксперимента по классификации текстовых отзывов по тональности

Для апробации алгоритмов определения тональности сформирована база из 5037 комментариев с сайта *prodoctorov.ru* с наличием исходной разметки по тональности и объекту применения.

В качестве алгоритма векторизации текстовых данных использовалась языковая модель *RuBERT*. Для бинарной классификации текста на категории (положительный или отрицательный) использовалась модель Трансформер. Соотношение обучающей и тестовой выборок составило 80/20. Результаты работы классификатора на тестовой выборке: *Precision* = 0,9857, *Recall* = 0,8909, *F1-score* = 0,9359. Полученные значения метрик качества работы классификатора позволяют утверждать о возможности применения данной архитектуры бинарного классификатора по тональности текста на источники данных медицинских отзывов.

На данной выборке также апробирован *LSTM*-классификатор (см. подраздел 1.1). Соотношение обучающей и тестовой выборок составило 80/20. Результаты классификации отзывов с сайта *prodoctorov.ru* с применением *LSTM*-сети следующие: позитивный отзыв о враче – 21%; позитивный отзыв о клинике – 54%; негативный отзыв о враче – 5%; негативный отзыв о клинике – 20%.

3.3 Эксперимент по классификации текстовых отзывов с применением различных моделей машинного обучения

Для классификации отзывов с использованием представленных в разделе 1 моделей машинного обучения использовались данные агрегатора *infodoctor.ru*. Преимущество этого агрегатора перед площадками (*prodoctorov.ru*, *docdoc.ru*) состоит в наличии группировки отзывов по рейтингу от одного до пяти для разных городов, что упрощает процедуру сбора данных. Выборки охватывают Москву, Санкт-Петербург и 14 других городов-миллионников России, по которым есть возможность сформировать минимально представительные выбор-

ки (не менее 1000 наблюдений в расчёте на один город), в период с июля 2012 года по август 2023 года. Всего извлечено 58246 отзывов. Примеры выбранных отзывов приведены в таблице 1 (тексты отзывов даны в оригинальном виде).

Таблица 1 – Примеры отзывов с сайта *infodoctor.ru*

№	Текст отзыва	Информация об отзыве	Класс тональности	Класс объекта применения
1	«Врач очень грубая, совершенно не имеет подход к людям, ваше заболевание её не интересует, для неё важнее пораньше уйти домой. Обращаться к такому врачу больше не хочется. Никому её не советую»	Екатерина, 13.04.2023 г., г. Москва	Отрицательный	О враче
2	«Проходил МРТ брюшной полости. Приняли меня не вовремя. Результаты по обследованию мне выдали сразу, обращаюсь с ними к врачу. Добираться до клиники мне было удобно. Общались со мной не очень хорошо. Не приду сюда повторно»	Камиль, 17.04.2023 г., г. Москва	Отрицательный	О клинике
3	«Все положительные отзывы составлены маркетологами, с отрицательными они всюду тщательно борются, реальные негативный отзывы пресекают. Клиника очень дорогая и ориентирована чисто на сбор денег, на ваше здоровье там всем глубоко фиолетово»	Анонимный пользователь, 10.04.2023 г., г. Москва	Отрицательный	О клинике
4	«В этой клинике происходит сбор денег, потому что назначают ненужно обследование и ненужные анализы - это я узнала уже, когда все начала проходить, а потом задумалась, зачем мне все это»	Арина, 2.03.2023 г., г. Москва	Отрицательный	О клинике
5	«Плохой врач. Моя проблема - сильная сухость кожи и высыпание на этой почве. ##### сказала только «её надо увлажнять» и всё. Как и чем не сказала. Рекомендации по уходу или дальнейшим действиям пришлось выпрашивать. Ничего не назначила кроме одного крема, после просьб моих»	Без имени, 11.05.2023 г., г. Москва	Отрицательный	О враче

Соотношение обучающей и тестовой выборок для всех применяемых алгоритмов составило 80/20. Графики, отражающие результаты классификации на обучающем и тестовом наборах данных для *LSTM*-, *GRU*- и *CNN*-архитектурах, представлены на рисунке 5.

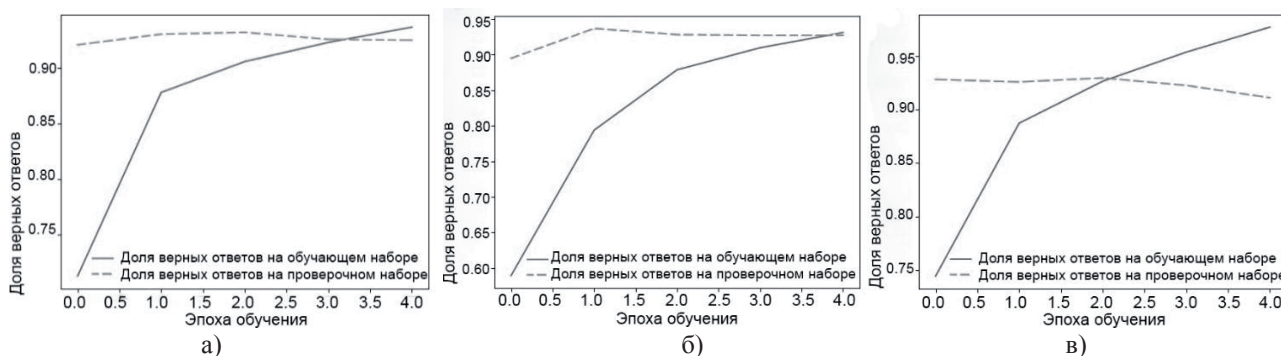


Рисунок 5 - Результаты классификации на обучающем и тестовом наборах данных для *LSTM*-сети (а), *GRU*-сети (б) и *CNN*-сети (в).

Сравнение показателей эффективности классификации текстовых отзывов представлено в таблице 2, где: *Accuracy* – точность при обучении; *Val_accuracy* – точность при проверке; *Loss* – потери при обучении; *Val_loss* – потери при проверке. Для оценки достоверности показателей характеристик предложенных моделей на используемом наборе данных проведены эксперименты с применением *SVM* и *RuBERT*. Из таблицы 2 видно, что эти алгоритмы показали чуть меньшую эффективность.

Одной из особенностей анализируемых текстовых отзывов стало наличие элементов разных классов внутри одного отзыва: одно текстовое сообщение могло содержать отзыв о враче и о клинике. В связи с этим введено два класса – положительный и отрицательный отзывы о клинике и враче – и применён лингвистический метод (*Ling*) поиска именованных сущностей (см. подраздел 1.4). Применение данного подхода позволило повысить эффективность классификации для всех трёх архитектур искусственных НС. Результаты классификации с использованием рассмотренных алгоритмов представлены на рисунке 6.

Лингвистический подход применялся только к тем сообщениям, которые на первом этапе НС отнесла к категории «отзыв о клинике».

Применение метода поиска именованных сущностей к текстовым сообщениям позволило повысить эффективность классификации.

Среди отзывов, неправильно классифицированных (в т.ч. с учётом применения *Ling*), имеется набор длинных текстовых сообщений, которые семантически могут относиться одновременно к разным классам. Это отзывы, которые:

- характеризуют клинику и врача, но без упоминания конкретного имени, что не позволяет отнести отзыв к смешанному классу;
- включают противоположные высказывания о клинике, которые касаются разных сторон её функционирования (в первую очередь могут различаться мнения об организационном сопровождении и уровне медицинского обслуживания клиник). Более глубокая категоризация отзывов о клинике позволит повысить качество классификации.

Заключение

В статье представлен метод классификации текстовых отзывов о работе медицинских учреждений, извлечённых из социальных медиаресурсов. Данный метод предполагает применение одной из архитектур искусственных НС (*LSTM*, *CNN*, *GRU*) для классификации отзывов на основные категории (по тональности и по объекту обращения) и лингвистического подхода с извлечением именованных сущностей.

Проведён ряд экспериментов по классификации текстовых отзывов о медицинских услугах клиник и врачей. В качестве исходных данных использовались текстовые отзывы агрегаторов *prodoctorov.ru* и *infodoctor.ru*. Проанализировано около 60 тысяч отзывов. По результатам проведённых экспериментов можно сделать следующие выводы.

- Классификация русскоязычных текстовых отзывов, извлечённых из социальных медиа, о медицинских учреждениях и врачах с использованием НС, имеет высокую эффектив-

Таблица 2 – Показатели эффективности классификации текстовых отзывов

Показатели	<i>LSTM</i>	<i>GRU</i>	<i>CNN</i>	<i>SVM</i>	<i>BERT</i>
<i>Accuracy</i>	0.9369	0.9309	0.9772	0.8441	0.8942
<i>Val accuracy</i>	0.9253	0.9271	0.9112	0.8289	0.8711
<i>Loss</i>	0.1859	0.2039	0.0785	0.3769	0.1729
<i>Val loss</i>	0.2248	0.2253	0.3101	0.3867	0.2266

Применение данного подхода позволило повысить **Val_accuracy**

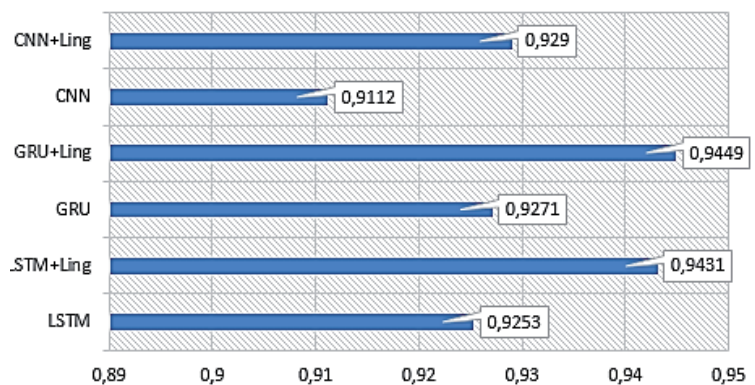


Рисунок 6 - Результаты классификации с использованием рассмотренных алгоритмов

ность при разбиении по тональности и объекту приложения. Наиболее высокую эффективность классификации показала архитектура на основе *GRU* (*val_accuracy*=0.9271).

- Применение метода поиска именованных сущностей к текстовым сообщениям позволило повысить эффективность классификации.

Авторский вклад

Калабихина И.Е. - постановка и разработка задачи исследования; Мошкин В.С. - разработка и формализация моделей и алгоритмов; Колотуша А.В. - сбор и предобработка обучающего и тестового множества; Кашин М.И. - программная реализация программной системы; Клименко Г.А. - анализ современных подходов; Казбекова З.Г. - разработка и оценка применимости алгоритмов, подготовка обучающего и тестового множества.

Список источников

- [1] *Litvin S.W., Goldsmith R.E., Pan B.* Electronic word-of-mouth in hospitality and tourism management. *Tourism management*. 2008; 29(3), 458-468. DOI: 10.1016/j.tourman.2007.05.011.
- [2] *Emmert M., Meier F.* An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. *Journal of medical Internet research*. 2013; 15(8), e2655. DOI:10.2196/jmir.2655.
- [3] *Nwachukwu B.U., Adjei J., Trehan S.K., Chang B., Amoo-Achampong K., Nguyen J.T., Ranawat A.S.* Rating a sports medicine surgeon's "quality" in the modern era: an analysis of popular physician online rating websites. *HSS Journal*. 2016; 12(3), 272-277. DOI: 10.1007/s11420-016-9520-x.
- [4] *Obele C.C., Duszak Jr.R., Hawkins C.M., Rosenkrantz A.B.* What patients think about their interventional radiologists: assessment using a leading physician ratings website. *Journal of the American College of Radiology*. 2017; 14(5), 609-614. DOI: 10.1016/j.jacr.2016.10.013.
- [5] *Emmert M., Meier F., Pisch F., Sander U.* Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. *Journal of medical Internet research*. 2013; 15(8), e2702. DOI: 10.2196/jmir.2702.
- [6] *Gao G.G., McCullough J.S., Agarwal R., Jha A.K.* A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *Journal of medical Internet research*. 2012; 14(1), e38. DOI: 10.2196/jmir.2003.
- [7] *Galizzi M.M., Miraldo M., Stavropoulou C., Desai M., Jayatunga W., Joshi M., Parikh S.* Who is more likely to use doctor-rating websites, and why? A cross-sectional study in London. *BMJ open*. 2012; 2(6), e001493. DOI: 10.1136/bmjopen-2012-001493.
- [8] *Hanauer D.A., Zheng K., Singer D.C., Gebremariam A., Davis M.M.* Public awareness, perception, and use of online physician rating sites. *Jama*. 2014; 311(7), 734-735. DOI: 10.1001/jama.2013.283194.
- [9] *McLennan S., Strech D., Meyer A., Kahrass H.* Public awareness and use of German physician ratings websites: Cross-sectional survey of four North German cities. *Journal of medical Internet research*. 2017; 19(11), e387. DOI: 10.2196/jmir.7581.
- [10] *Lin Y., Hong Y.A., Henson B.S., Stevenson R.D., Hong S., Lyu T., Liang C.* Assessing patient experience and healthcare quality of dental care using patient online reviews in the United States: mixed methods study. *Journal of Medical Internet Research*. 2020; 22(7), e18652. DOI: 10.2196/18652.
- [11] *Emmert M., Meier F., Heider A.K., Dürr C., Sander U.* What do patients say about their physicians? An analysis of 3000 narrative comments posted on a German physician rating website. *Health policy*. 2014; 118(1), 66-73. DOI: 10.1016/j.healthpol.2014.04.015.
- [12] *Shah A.M., Yan X., Shah S.A.A., Mamirkulova G.* Mining patient opinion to evaluate the service quality in healthcare: a deep-learning approach. *Journal of Ambient Intelligence and Humanized Computing*. 2020; 11, 2925-2942. DOI: 10.1007/S12652-019-01434-8.
- [13] *Jiang S., Street R.L.* Pathway linking internet health information seeking to better health: a moderated mediation study. *Health Communication*. 2017; 32(8), 1024-1031. DOI: 10.1080/10410236.2016.1196514.
- [14] *Hotho., Nürnberger A., Paaß G.* A Brief Survey of Text Mining, *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*. 2005; vol. 20, pp.19-62. DOI: 10.21248/jlcl.20.2005.68.
- [15] *Păvăloaia V., Teodor E., Fotache D., Danileț M.* Opinion Mining on Social Media Data: Sentiment Analysis of User Preferences, *Sustainability*. 2019; 11, 4459. DOI: 10.3390/su11164459.
- [16] *Bespalov D., Bing B., Yanjun Q., Shokoufandeh A.* Sentiment classification based on supervised latent n-gram analysis", *Proceedings of the 20th ACM international conference on Information and knowledge management*

- (CIKM '11). Association for Computing Machinery. 2011; New York, USA, 375–382. DOI: 10.1145/2063576.2063635.
- [17] **Irfan R, King CK, Grages D, Ewen S, Khan SU, Madani SA, Kolodziej J, Wang L, Chen D, Rayes A, Tziritas N, Xu CZ, Zomaya AY, Alzahrani AS, Li H.** A Survey on Text Mining in Social Networks, Cambridge Journal, The Knowledge Engineering Review. 2015; 30(2), pp. 157-170. DOI:10.1017/S0269888914000277.
- [18] **Patel P., Mistry K.** A Review: Text Classification on Social Media Data, IOSR Journal of Computer Engineering. 2015; 17(1), pp. 80-84.
- [19] **Lee K., Palseia D., Narayanan R., Patwary Md.M.A, Agrawal A., Choudhary A.S.** Twitter Trending Topic Classification, in Proceeding of the 2011 IEEE 11 th International Conference on Data Mining Workshops, ICDW'11. 2011; pp. 251-258.
- [20] **Kateb F., Kalita J.** Classifying Short Text in Social Media: Twitter as Case Study, International Journal of Computer Applications. 2015; 111(9), pp. 1-12. DOI: 10.5120/19563-1321.
- [21] **Theeramunkong T., Lertnattee V.** Multi-Dimension Text Classification, SIIT, Thammasat University, 2005. <http://www.aclweb.org/anthology/C02-1155> (03 October 2024). DOI: 10.1109/TITB.2004.832.
- [22] **Sornlertlamvanich V., Pacharawongsakda E., Charoenporn T.** Understanding Social Movement by Tracking the Keyword in Social Media, in MAPLEX2015, 2015; Yamagata, Japan.
- [23] **Kalabikhina I., Moshkin V., Kolotusha A., Kashin M., Klimenko G., Kazbekova Z.** Advancing Semantic Classification: A Comprehensive Examination of Machine Learning Techniques in Analyzing Russian-Language Patient Reviews. Mathematics. 2024; 12(4): 566. DOI:10.3390/math12040566.
- [24] **Kalabikhina I., Zubova E., Loukachevitch N., Kolotusha A., Kazbekova Z., Banin E., Klimenko G.** Identifying Reproductive Behavior Arguments in Social Media Content Users' Opinions through Natural Language Processing Techniques, Population and Economics. 2023; 7(2), pp. 40-59. DOI: 10.3897/popecon.7.e97064.

Сведения об авторах



Калабихина Ирина Евгеньевна, 1969 г. рождения. Доктор экономических наук, профессор. Заведующая кафедрой народонаселения Экономического факультета МГУ. Главный редактор журнала «Население и экономика». Области исследований: взаимосвязь демографического и экономического развития, прикладная демография, экономика неравенства, детерминанты рождаемости, социальное бюджетирование, политика в области народонаселения, политика в интересах социально-демографических групп населения. Author ID (Scopus): 57190138890; Researcher ID (WoS): N-3625-2013; ORCID: 0000-0002-3958-6630. ikalabikhina@yandex.ru.

Мошкин Вадим Сергеевич, 1990 г. рождения. Окончил УлГТУ в 2012 г., к.т.н. (2017 г.), доцент кафедры «Информационные системы» УлГТУ. Проректор по цифровой трансформации УлГТУ. Вице-президент Российской ассоциации искусственного интеллекта. В списке научных трудов более 150 статей в области интеллектуальной обработки знаний, автоматизации проектирования, построения прикладных интеллектуальных систем. Author ID (RSCI): 762084; Author ID (Scopus): 57190250573; Researcher ID (WoS): L-3578-2016; ORCID: 0000-0002-9258-4909. v.moshkin@ulstu.ru ✉.



Колотуша Антон Васильевич, 1995 г. рождения. Кандидат экономических наук. Программист 2 категории лаборатории информационно-аналитических ресурсов кафедры народонаселения Экономического факультета МГУ. Области исследований: прикладная демография, экономика неравенства, анализ социальных медиа, политика в области народонаселения. Author ID (Scopus): 57223390271; ORCID: 0000-0002-3089-3327. tony_kol@mail.ru.

Кашин Максим Игоревич, 2003 г. рождения. Студент направления 09.03.04 «Программная инженерия» УлГТУ. Младший научный сотрудник научно-исследовательского отдела Департамента научных исследований и инноваций УлГТУ. Области исследований: анализ естественного языка, принятие решений, машинное обучение, автоматизация сложных бизнес-процессов. m.kashin@ulstu.ru.



Клименко Герман Андреевич. Аспирант кафедры народонаселения Экономического факультета МГУ. Области исследований: прикладная демография, экономика неравенства, анализ социальных медиа, политика в области народонаселения. Author ID (Scopus): 57223391224; ORCID: 0000-0003-3045-7151. german89000@mail.ru.



Казбекова Зарина Германовна. Кандидат экономических наук. Научный сотрудник кафедры народонаселения Экономического факультета МГУ. Выпускник аспирантуры кафедры народонаселения. Области исследований: прикладная демография, экономика неравенства, анализ социальных медиа, политика в области народонаселения. Author ID (Scopus): 57934120000; ORCID: 0000-0002-7567-3184. kazbekova.zarina@bk.ru.



Поступила в редакцию 5.09.2024, после рецензирования 2.12.2024. Принята к публикации 18.12.2024.



Scientific article

DOI: 10.18287/2223-9537-2025-15-1-55-66

Analysis of patient reviews using machine learning and linguistic methods

© 2025, I.E. Kalabikhina¹, V.S. Moshkin²✉, A.V. Kolotusha¹, M.I. Kashin², G.A. Klimenko¹, Z.G. Kazbekova¹

¹Moscow State University named after M.V. Lomonosov, Moscow, Russia

²Ulyanovsk State Technical University, Ulyanovsk, Russia

Abstract

With the advancement of digitalization, traditional methods of surveying consumers to assess their satisfaction with service quality are being replaced by approaches based on the automatic processing of text data from social media. This study aims to determine the degree of patient satisfaction with the quality of medical services by developing and testing an algorithm for classifying Russian-language text reviews collected from social media platforms. The focus is on analyzing the sentiment (positive/negative) of patient reviews about medical institutions and doctors, as well as identifying the review's subject—either the quality of medical services provided or the organization of patient care by the institution. A method was developed for classifying text reviews about the work of medical institutions posted by patients on two Russian doctor review platforms. Approximately 60,000 reviews were analyzed. Machine learning techniques, including various artificial neural network architectures, were tested. The classification algorithm demonstrated high efficiency, with the best performance achieved using a recurrent neural network architecture (accuracy = 0.9271). Incorporating named entity recognition into text analysis further enhanced the classification efficiency across all neural network-based classifiers. To improve classification quality, the study highlights the need for semantic segmentation of reviews by their subject and sentiment, followed by the separate analysis of these fragments.

Keywords: machine learning, patient reviews, neural networks, review classification, quality of medical services.

For citation: Kalabikhina IE, Moshkin VS, Kolotusha AV, Kashin MI, Klimenko GA, Kazbekova ZG. Analysis of patient reviews using machine learning and linguistic methods [In Russian]. *Ontology of designing*. 2025; 15(1): 55-66. DOI:10.18287/2223-9537-2025-15-1-55-66.

Financial Support: This work was supported by the Russian Science Foundation, project No. 23-71-01101 “Development of models and methods for increasing the performance of data warehouses through predictive analysis of temporal diagnostic information» and research work “Population reproduction in socio-economic development” with the support of the Faculty of Economics of the Lomonosov Moscow State University.

Conflict of interest: The authors declare no conflict of interest.

List of figures and table

Figure 1 - LSTM network: general architecture, LSTM layer

Figure 2 - General architecture of a recurrent neural network

Figure 3 - General architecture of a convolutional neural network

Figure 4 - General scheme of the classification algorithm

Figure 5 - Results of classification on the training and test dataset for an LSTM network, a GRU network, and a CNN network

Figure 6 - Classification results using the considered algorithms

Table 1 - Examples of reviews from the infodoctor.ru website

Table 2 - Efficiency indicators for classifying text reviews

References

- [1] **Litvin SW, Goldsmith RE, Pan B.** Electronic word-of-mouth in hospitality and tourism management. *Tourism management*. 2008; 29(3): 458-468. DOI: 10.1016/j.tourman.2007.05.011.
- [2] **Emmert M, Meier F.** An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. *Journal of medical Internet research*. 2013; 15(8): e2655. DOI:10.2196/jmir.2655.
- [3] **Nwachukwu BU, Adjei J, Trehan SK, Chang B, Amoo-Achampong K, Nguyen JT, Ranawat AS.** Rating a sports medicine surgeon's "quality" in the modern era: an analysis of popular physician online rating websites. *HSS Journal*. 2016; 12(3): 272-277. DOI: 10.1007/s11420-016-9520-x.
- [4] **Obele CC, Duszak JrR, Hawkins CM, Rosenkrantz AB.** What patients think about their interventional radiologists: assessment using a leading physician ratings website. *Journal of the American College of Radiology*. 2017; 14(5): 609-614. DOI: 10.1016/j.jacr.2016.10.013.
- [5] **Emmert M, Meier F, Pisch F, Sander U.** Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. *Journal of medical Internet research*. 2013; 15(8): e2702. DOI: 10.2196/jmir.2702.
- [6] **Gao GG, McCullough JS, Agarwal R, Jha AK.** A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *Journal of medical Internet research*. 2012; 14(1): e38. DOI: 10.2196/jmir.2003.
- [7] **Galizzi MM, Miraldo M, Stavropoulou C, Desai M, Jayatunga W, Joshi M, Parikh S.** Who is more likely to use doctor-rating websites, and why? A cross-sectional study in London. *BMJ open*. 2012; 2(6): e001493. DOI: 10.1136/bmjopen-2012-001493.
- [8] **Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM.** Public awareness, perception, and use of online physician rating sites. *Jama*. 2014; 311(7): 734-735. DOI: 10.1001/jama.2013.283194.
- [9] **McLennan S, Strech D, Meyer A, Kahrass H.** Public awareness and use of German physician ratings websites: Cross-sectional survey of four North German cities. *Journal of medical Internet research*. 2017; 19(11): e387. DOI: 10.2196/jmir.7581.
- [10] **Lin Y, Hong YA, Henson BS, Stevenson RD, Hong S, Lyu T, Liang C.** Assessing patient experience and healthcare quality of dental care using patient online reviews in the United States: mixed methods study. *Journal of Medical Internet Research*. 2020; 22(7): e18652. DOI: 10.2196/18652.
- [11] **Emmert M, Meier F, Heider AK, Dürr C, Sander U.** What do patients say about their physicians? An analysis of 3000 narrative comments posted on a German physician rating website. *Health policy*. 2014; 118(1): 66-73. DOI: 10.1016/j.healthpol.2014.04.015.
- [12] **Shah AM, Yan X, Shah SAA, Mamirkulova G.** Mining patient opinion to evaluate the service quality in healthcare: a deep-learning approach. *Journal of Ambient Intelligence and Humanized Computing*. 2020; 11: 2925-2942. DOI: 10.1007/S12652-019-01434-8.
- [13] **Jiang S, Street RL.** Pathway linking internet health information seeking to better health: a moderated mediation study. *Health Communication*. 2017; 32(8): 1024-1031. DOI: 10.1080/10410236.2016.1196514.
- [14] **Hotho, Nürnberger A, Paaf G.** A Brief Survey of Text Mining, LDV Forum - GLDV *Journal for Computational Linguistics and Language Technology*. 2005; 20: 19-62. DOI: 10.21248/jlcl.20.2005.68.
- [15] **Păvăloaia V, Teodor E, Fotache D, Danileț M.** Opinion Mining on Social Media Data: Sentiment Analysis of User Preferences, *Sustainability*. 2019; 11: 4459. DOI: 10.3390/su11164459.
- [16] **Bespalov D, Bing B, Yanjun Q, Shokoufandeh A.** Sentiment classification based on supervised latent n-gram analysis". Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM '11). Association for Computing Machinery. 2011; NY, USA, 375-382. DOI: 10.1145/2063576.2063635.
- [17] **Irfan R, King CK, Grages D, Ewen S, Khan SU, Madani SA, Kolodziej J, Wang L, Chen D, Rayes A, Tziritas N, Xu CZ, Zomaya AY, Alzahrani AS, Li H.** A Survey on Text Mining in Social Networks, Cambridge Journal, *The Knowledge Engineering Review*. 2015; 30(2): 157-170. DOI:10.1017/S0269888914000277.
- [18] **Patel P, Mistry K.** A Review: Text Classification on Social Media Data, *IOSR Journal of Computer Engineering*. 2015; 17(1): 80-84.

- [19] **Lee K, Palsetia D, Narayanan R, Patwary MdMA, Agrawal A, Choudhary AS.** Twitter Trending Topic Classification. In Proceeding of the 2011 IEEE 11 th International Conference on Data Mining Workshops, ICDW'11. 2011; pp. 251-258.
- [20] **Kateb F, Kalita J.** Classifying Short Text in Social Media: Twitter as Case Study. *International Journal of Computer Applications*. 2015; 111(9): 1-12. DOI: 10.5120/19563-1321.
- [21] **Theeramunkong T, Lertnattee V.** Multi-Dimension Text Classification, SIIT, Thammasat University, 2005. <http://www.aclweb.org/anthology/C02-1155> (03 October 2024). DOI: 10.1109/TITB.2004.832.
- [22] **Sornlertlamvanich V, Pacharawongsakda E, Charoenporn T.** Understanding Social Movement by Tracking the Keyword in Social Media, in MAPLEX2015, 2015; Yamagata, Japan.
- [23] **Kalabikhina I, Moshkin V, Kolotusha A, Kashin M, Klimenko G, Kazbekova Z.** Advancing Semantic Classification: A Comprehensive Examination of Machine Learning Techniques in Analyzing Russian-Language Patient Reviews. *Mathematics*. 2024; 12(4):566. DOI:10.3390/math12040566.
- [24] **Kalabikhina I, Zubova E, Loukachevitch N, Kolotusha A, Kazbekova Z., Banin E, Klimenko G.** Identifying Reproductive Behavior Arguments in Social Media Content Users' Opinions through Natural Language Processing Techniques, *Population and Economics*. 2023; 7(2): 40-59. DOI: 10.3897/popecon.7.e97064.
-

About the authors

Irina Evgenievna Kalabikhina (b. 1969). Doctor of Economics, Professor. Head of the Population Department of the Faculty of Economics of Moscow State University. Editor-in-Chief of the journal "Population and Economics". Research areas are the relationship between demographic and economic development, applied demography, economics of inequality, determinants of fertility, social budgeting, population policy, policy in the interests of socio-demographic groups of the population. Author ID (Scopus): 57190138890; Researcher ID (WoS): N-3625-2013; ORCID: 0000-0002-3958-6630. ikalabikhina@yandex.ru.

Vadim Sergeevich Moshkin (b. 1990). Graduated from UISTU in 2012, PhD (2017), Associate Professor of the Information Systems Department of UISTU. Vice-Rector for Digital Transformation of UISTU. Vice-President of the Russian Association of Artificial Intelligence. The list of scientific works includes more than 150 articles in the field of intelligent knowledge processing, design automation, and the construction of applied intelligent systems. Author ID (RSCI): 762084; Author ID (Scopus): 57190250573; Researcher ID (WoS): L-3578-2016; ORCID: 0000-0002-9258-4909. v.moshkin@ulstu.ru ✉.

Anton Vasilievich Kolotusha (b. 1995). Candidate of Economic Sciences. 2nd category programmer of the of the Laboratory of Information and Analytical Resources of the Department of Population of the Economics Faculty of Moscow State University named after M.V. Lomonosov. The areas of research include applied demography, economics of inequality, social media analysis, population policy. Author ID (Scopus): 57223390271; ORCID: 0000-0002-3089-3327. tony_kol@mail.ru.

Maksim Igorevich Kashin (b. 2003). Student of the "Software Engineering" specialty of the Ulyanovsk State Technical University (UISTU). Junior researcher of the Research Department of the Department of Scientific Research and Innovation of UISTU. The areas of research include natural language analysis, decision making, machine learning, automation of complex business processes. m.kashin@ulstu.ru.

German Andreevich Klimenko. Postgraduate student of the Department of Population of the Economics Faculty of Moscow State University. The areas of research include applied demography, economics of inequality, social media analysis, and population policy. Author ID (Scopus): 57223391224; ORCID: 0000-0003-3045-7151. german89000@mail.ru.

Zarina Germanovna Kazbekova. Candidate of Economic Sciences. Researcher of the Department of Population of the Economics Faculty of Moscow State University. Graduate of the postgraduate program of the Department of Population. The areas of research include applied demography, economics of inequality, social media analysis, and population policy. Author ID (Scopus): 57934120000; ORCID: 0000-0002-7567-3184. kazbekova.zarina@bk.ru.

Received September 5, 2024. Revised December 2, 2024. Accepted December 18, 2024.
