

# Методика формирования набора признаков с целью описания объектов социальной сети

Ю. В. Катенко

Санкт-Петербургский государственный электротехнический университет

«ЛЭТИ» им. В.И. Ульянова (Ленина);

ООО «БалтИнфоКом»

katenkoyuliya@gmail.com

**Аннотация.** В работе рассматривается подход к формированию набора признаков, описывающих объекты социальной сети – пользователей, записи – и связи между ними. Представление объектов в виде вектора значений выбранных признаков дает возможность применить методы машинного обучения для анализа этих объектов.

**Ключевые слова:** выбор признаков, извлечение признаков, описание объектов анализа

## I. ВВЕДЕНИЕ

В наше время все больше людей получает новости из интернета и в том числе из социальных сетей, а также формирует мнение по различным вопросам, основываясь на найденной в сети информации. Поэтому важной проблемой является распространение в социальных сетях недостоверной, ложной, мошеннической информации. В этой работе рассматривается подход к выбору и вычислению набора признаков объектов социальной сети – записей и пользователей.

Признаком называется величина, описывающая объект. Признаки делятся на два типа:

- количественные — значения принадлежат некоторому подмножеству действительных чисел;
- категориальные — принимают конечное число значений и подразделяются на упорядочиваемые (например, значения «низкий», «средний», «высокий») и номинальные (тема записи, профессия пользователя и т. п.). [1]

Создание векторов признаков объектов является важным этапом построения системы оценки достоверности записей в социальной сети. Они необходимы для применения машинного обучения в целях обнаружения недостоверной информации.

## II. СВЯЗАННЫЕ РАБОТЫ

Выбору признаков, описывающих запись в социальной сети, посвящена статья O. Enayet и S. R. El-Beltagy «Rumour and Veracity Support for Rumours on Twitter». Авторы предлагают преобразовывать текст записи – токенизировать, удалять стоп-слова, знаки пунктуации и слова, присущие конкретной социальной сети (например, слова «rt» и «via» для Твиттера). В результате набор

признаков, описывающих текст, представляет собой мешок слов. Также предлагается ряд дополнительных признаков записи, таких как наличие в тексте вопроса, отрицательных и утвердительных слов, хештегов и ссылок, тональность записи и др. Кроме этого в набор признаков включается информация об авторе записи – количество его друзей и подписчиков, число записей на его странице, наличие у аккаунта статуса подтвержденного. Авторы рассматриваемой статьи использовали предложенные признаки в качестве входных данных для обучения моделей классификации записей по их достоверности, наибольшее значение точности классификации – 0.53 – было получено для линейного метода опорных векторов. [2] Можно сделать вывод, что приведенный в статье набор признаков не является достаточным и требует доработки.

В настоящее время большую популярность среди исследователей получили фальшивые новости (fake news), и, хотя записи в социальной сети нельзя полностью отождествлять с новостями, работы на эту тему представляют большой интерес. Так статья Г. Кузьмина и др. «Fake news detection for the Russian language» посвящена обнаружению недостоверных новостей на русском языке. В этой работе рассмотрены отдельно три класса текстов: достоверные новости, недостоверные новости и сатирические новости. Авторы экспериментально выяснили, что сатирические новости заметно отличаются от действительно фальшивых новостей. Несмотря на то, что люди часто принимают и те, и другие за достоверные, классификаторы показали лучший результат на наборе данных, где сатирические и фальшивые новости рассматриваются отдельно, а не как единый класс. Также интересен результат оценки достоверности новостей людьми (не экспертами) – несколько человек должны были отнести предложенные новости к одному из трех классов, и все испытуемые справились хуже, чем модель классификации. [3] Из этого можно сделать вывод, что создание качественного набора данных для обучения модели оценки достоверности текстов – весьма сложная задача.

## III. ВЫБОР ПРИЗНАКОВ

Социальная сеть может быть представлена в виде ориентированного графа, состоящего из вершин классов

Пользователь, Сообщество, Запись, Комментарий. Типы связей между вершинами приведены в табл. 1.

ТАБЛИЦА I Типы связей в графе социальной сети

Тип связи	Класс начальной вершины	Класс конечной вершины
Дружит с	Пользователь	Пользователь
Участник	Пользователь	Сообщество
Опубликовал	Пользователь, Сообщество	Запись, Комментарий
Оценил положительно	Пользователь	Запись, Комментарий
Оценил отрицательно	Пользователь	Запись, Комментарий
Комментарий к	Комментарий	Запись, Комментарий
Репост	Запись	Запись

Графовое представление дает возможность рассматривать в качестве признаков объектов свойства вершин, соответствующих этим объектам. Основным из таких свойств является центральность – мера, показывающая важность вершины графа. Ниже приведены показатели, позволяющие определить центральность вершины в графе.

Степень связности – число ребер, связанных с этой вершиной.

Степень близости – мера, которая показывает, насколько близко вершина находится к другим вершинам графа. Вычисляется как среднее значение длины кратчайшего пути ко всем другим вершинам.

Степень посредничества – мера, показывающая сколько кратчайших путей на графе проходит через вершину. Вычисляется следующим образом: для каждой пары вершин находятся все кратчайшие пути, и определяется доля путей, включающих выбранную вершину. Операция повторяется для всех пар вершин, результат суммируется.

Степень влияния – мера центральности вершины, основанная на важности ее соседей: чем больше у вершины соседних узлов с высокой степенью влияния, тем выше ее собственная степень влияния.

Также в набор признаков может быть включено такое свойство вершины как локальный коэффициент кластеризации – мера, показывающая, как хорошо связаны вершины, соседние с выбранной. Этот показатель вычисляется как отношение количества ребер между соседними вершинами к общему количеству возможных ребер между ними. [4]

Из текста записи можно извлечь дополнительные признаки с помощью методов обработки естественного языка, такие как тематика текста, его тональность и, собственно, оценка достоверности текста.

### 1. Тематика текста

Существует ряд тем, для которых распространение недостоверной информации может нанести наибольший вред обществу, например, темы медицины и здоровья,

экономики, безопасности. То, что текст относится к конкретной тематике, не может указывать на его достоверность или недостоверность, но позволяет выделить наиболее важные записи, оценить которые нужно в первую очередь. Тематика может определяться путем поиска в тексте ключевых слов. Другой способ – использование моделей классификации текстов.

В социальных сетях часто встречаются мошеннические, псевдонаучные или посвященные теориям заговоров тексты, которые практически всегда могут быть отнесены к недостоверным. Их также следует учитывать при обучении тематических классификаторов.

### 2. Тональность текста

Тональность текста – это выраженное в тексте эмоциональное мнение относительно некоторого объекта (объектов). Тональность можно измерять при помощи шкалы в диапазоне от положительного отношения до отрицательного. [5]

В рамках этой работы был создан классификатор для определения тональности текста. В качестве модели использовалась сверточная нейронная сеть, архитектура которой описана в статье [6], для приведения текстов к векторному виду применялась заранее подготовленная модель word2vec. Для обучения использовался набор размеченных текстов с новостных сайтов. [7] Оценка полученной модели представлена в табл. 2. В перспективе для дополнительного обучения классификатора планируется использовать набор размеченных коротких текстов, описанный в [8].

ТАБЛИЦА II Оценка модели анализа тональности текстов

	Обучающая выборка			Проверочная выборка		
Точность (accuracy)	0,8529			0,8149		
Точность (precision)	0,8602			0,8193		
Полнота	0,8529			0,8149		
F <sub>1</sub> -мера	0,8626			0,8252		
Матрица несоответствий	0	1		0	1	
	2160	603	0	233	75	0
	210	2553	1	39	269	1

### 3. Оценка достоверности текста

Наиболее сложным является построение этого классификатора, потому что при создании набора данных для обучения необходимо разметить тексты как достоверные и недостоверные. В качестве возможного источника недостоверных текстов можно рассматривать набор данных Fake news, содержащий подборку текстов ИА «Панорама» – сайта с юмористическими новостями, написанными по тем же правилам и шаблонам, что и настоящие новости. Этот набор данных входит в Taiga Corpus – корпус, содержащий тексты и связанную с ними мета-информацию, используемые для решения популярных задач машинного обучения. [9]

Однако, заметим, что новости и тексты в социальной сети в целом отличаются – новости имеют определенный стиль изложения, написаны литературным языком, в то время как записи пользователей часто содержат субъективную оценку и включают разговорную лексику. Также необходимо учитывать выводы о различии сатирических и действительно фальшивых новостей, сделанные в статье [3]. Следовательно, этого набора данных недостаточно для создания классификатора, необходимо провести сбор записей из социальных сетей и вручную оценить эти записи на достоверность, причем требуется, чтобы каждая запись была оценена несколькими людьми.

При сборе и разметке записей можно выбрать несколько источников с высоким уровнем достоверности (например, информационное агентство ТАСС и подобные ему крупные издания) и искать в них информацию, приведенную в записи. При обнаружении связанных новостей необходимо проверять, что запись им не противоречит, в таком случае она может считаться достоверной. Если же связанные новости не найдены, то информация в записи либо недостоверна, либо является частной, и ее правдивость нельзя проверить.

Признаки могут быть получены из связей записи с другими вершинами графа, представляющего социальную сеть. Можно выделить простые количественные признаки – число одобрительных и неодобрительных оценок записи, число комментариев к записи, доля комментариев с положительной/отрицательной тональностью.

Также интерес представляют свойства автора записи и пользователей, отреагировавших на нее. Помимо свойств, предложенных в [2], таких как возраст аккаунта, наличие статуса подтвержденного аккаунта, число друзей и т. п., в качестве важного признака следует рассматривать оценку достоверности высказываний пользователя. Этот параметр зависит от предыдущих действий пользователя – если он уже публиковал недостоверные записи или оставлял одобрительные комментарии к таким записям, оценка достоверности его высказываний должна быть отрицательной. Если же пользователь публиковал записи, помеченные как достоверные, или оставлял опровергающие комментарии к недостоверным записям, его оценка будет положительной.

#### IV. ЗАКЛЮЧЕНИЕ

Предложенный набор признаков можно использовать не только для оценки достоверности записей в социальной сети, но и для решения других задач, например, для поиска похожих пользователей или для оценки роли пользователя в группе.

Вычисление некоторых из выбранных признаков было программно реализовано в рамках системы анализа и визуализации данных в графовом виде «Октопус» и использовано для кластеризации участников сообществ в социальной сети. Другие признаки, в первую очередь оценка достоверности текста записи и оценка достоверности высказываний пользователя, требуют дополнительного исследования.

Также в перспективе планируется рассмотрение процесса распространения информации в социальной сети в терминах теории фракталов, что позволит дополнить описание объектов анализа.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Glossary of Terms. Special Issue on Applications of Machine Learning and the Knowledge Discovery Process // Stanford Artificial Intelligence Laboratory URL: <https://ai.stanford.edu/~ronnyk/glossary.html> (дата обращения: 13.03.2021).
- [2] Omar Enayet, Samhaa R. El-Beltagy. NileTMRG at SemEval-2017 Task 8: Determining Rumour and Veracity Support for Rumours on Twitter. URL: <https://www.aclweb.org/anthology/S17-2082.pdf> (дата обращения: 13.03.2021).
- [3] Gleb Kuzmin, Daniil Larionov, Dina Pisarevskaya, Ivan Smirnov. Fake news detection for the Russian language. URL: <https://www.aclweb.org/anthology/2020.rdsm-1.5.pdf> (дата обращения: 13.03.2021).
- [4] Analyzing the Social Web / Jennifer Golbeck. USA: Morgan Kaufmann, 2013.
- [5] Котельников Е.В. Комбинированный метод автоматического определения тональности текста // Программные продукты и системы. 2012. №3. URL: <https://cyberleninka.ru/article/n/kombinirovannyi-metod-avtomaticheskogo-opredeleniya-tonalnosti-teksta> (дата обращения: 10.03.2021).
- [6] Yoon Kim. Convolutional Neural Networks for Sentence Classification. URL: <https://arxiv.org/pdf/1408.5882.pdf> (дата обращения: 10.03.2021).
- [7] Sentiment Analysis in Russian. Determine sentiments (positive, negative or neutral) of news in russian language. // Kaggle URL: <https://www.kaggle.com/c/sentiment-analysis-in-russian> (дата обращения: 10.03.2021).
- [8] Рубцова Ю. Автоматическое построение и анализ корпуса коротких текстов (постов микроблогов) для задачи разработки и тренировки тонового классификатора // Инженерия знаний и технологии семантического веба. 2012. Т. 1. С. 109-116.
- [9] Taiga Corpus URL: [https://tatianashavrina.github.io/taiga\\_site](https://tatianashavrina.github.io/taiga_site) (дата обращения: 13.03.2021)