УДК 519-6;519-8
DOI: 10.26102/2310-6018/2024.45.2.016

Прогнозирование депрессии на основе пользовательских данных русскоязычной социальной сети

Т.Д. Солохов✉, А.А. Кочкаров

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Резюме. В статье показаны возможности применения семантического анализа постов пользователей социальной сети ВКонтакте для мониторинга и прогнозирования депрессии. Подчеркивается серьезность проблемы депрессии, ее негативное влияние на здоровье и социум, а также актуальность ранней диагностики и помощи. В работе также обоснована необходимость и перспективы анализа данных русскоязычных социальных сетей для предотвращения развития депрессии у пользователей. В статье предложен подход, который позволяет проводить анализ текстовых данных и использовать логистическую регрессию для классификации пользователей по наличию депрессии. Результаты исследования показывают высокую точность модели с использованием логистической регрессии, что представляет потенциал для автоматизации процессов выявления и поддержки пользователей, страдающих депрессией по данным пользовательской информации в социальных сетях. Также приводится значимость данного метода, его практическая полезность для персонализированных интервенций, преимущества и перспективы развития, включая использование нейросетевых методов и анализ динамики данных. Несмотря на достигнутые результаты, отмечается необходимость дальнейшей работы над моделью, включая изучение других методов машинного обучения и учет изменений в психическом состоянии пользователя со временем. Развитие методов прогнозирования депрессии на основе данных социальных сетей, предложенных в статье, является важным направлением, способным принести значительный вклад в области психологии, здравоохранения и информационных технологий.

Ключевые слова: прогнозирование, депрессия, психическое расстройство, логистическая регрессия, классификация, социальная сеть, машинное обучение.

Для цитирования: Солохов Т.Д., Кочкаров А.А. Прогнозирование депрессии на основе пользовательских данных русскоязычной социальной сети. *Моделирование, оптимизация и информационные технологии.* 2024;12(2). URL: https://moitvivt.ru/ru/journal/pdf?id=1593 DOI: 10.26102/2310-6018/2024.45.2.016

Forecasting the depression with user data from Russian-language social network

T.D. Solokhov✉, A.A. Kochkarov

Financial University under the Government of the Russian Federation, Moscow, the Russian Federation

Abstract. The article explores the possibilities of applying semantic analysis of user posts on the social network VKontakte for monitoring and predicting depression. It emphasizes the seriousness of the depression issue, its negative impact on health and society, and the relevance of early diagnosis and assistance. The study also justifies the necessity and prospects of analyzing data from Russian-language social networks to prevent the development of depression among users. The article examines the analysis of textual data and the use of logistic regression to classify users based on the presence of depression. The study's results show high model accuracy using logistic regression, demonstrating the potential for automating the processes of identifying and supporting users suffering from depression in the online

© Солохов Т.Д., Кочкаров А.А., 2024

environment based on user information from social networks. The significance of this method is also highlighted, along with its practical usefulness for personalized interventions, its advantages, and its development prospects, including the use of neural network methods and the analysis of data dynamics. Despite the results achieved, there is a need for further work on the model, including the study of other machine learning methods and taking into account changes in the user's mental state over time. The development of depression prediction methods based on social network data, as proposed in the article, is an important direction that can make a significant contribution to psychology, healthcare, and information technology.

Keywords: forecasting, depression, psychological disorder, logistic regression, classification, social network, machine learning.

For citation: Solokhov T.D., Kochkarov A.A. Forecasting the depression with user data from Russian-language social network. *Modeling, Optimization and Information Technology*. 2024;12(2). URL: https://moitvivt.ru/ru/journal/pdf?id=1593 DOI: 10.26102/2310-6018/2024.45.2.016 (In Russ.).

Введение

Депрессия – это психическое расстройство, которое сопровождается постоянным чувством печали, утраты интереса к жизни, апатией и ухудшением качества жизни[1]. Она оказывает значительное влияние на здоровье и благополучие человека, приводя к снижению производительности, социальной изоляции, а в некоторых случаях – к самоубийству. Несмотря на масштаб проблемы, депрессия часто остается недооцененной и не диагностированной.

Согласно исследованиям, около 280 миллионов человек в мире находятся в состоянии глубокой депрессии, что составляет около 4 % населения Земли[2].

Прогнозирование депрессии на ранних стадиях может сыграть ключевую роль в предотвращении ее развития и оказании своевременной помощи людям, подверженным этому состоянию. Помимо негативного воздействия непосредственно на самого человека, депрессия может оказывать прямое влияние и на окружающих. Приведем пример: находящийся в глубокой депрессии пилот Germanwings совершил суицид, устроив авиакатастрофу, принеся с собой в жертву еще 144 пассажира и 5 членов экипажа, при этом пилот активно вел социальные сети и жаловался на свое неудовлетворительное эмоциональное состояние[3].

Социальные сети предоставляют уникальную среду, в которой пользователи часто выражают свои эмоции, переживания и мысли, в том числе и относительно своего психического состояния. Семантический анализ постов пользователей в социальных сетях, таких как ВКонтакте (VK), может стать основой для мониторинга и прогнозирования депрессии.

Однако на данный момент есть лишь немногочисленные исследования, посвященные прогнозированию депрессии на основе русскоязычных социальных сетей, таких как ВКонтакте. Это оставляет пробел в понимании возможностей и ограничений данного метода в контексте российской аудитории и языковой среды.

Таким образом, целью исследования является апробация метода прогнозирования и мониторинга депрессии на основе семантического анализа публикаций пользователей социальной сети ВКонтакте, а также на основе данных профиля пользователя – его

[1] International Statistical Classification of Diseases and Related Health Problems 10th Revision. URL: https://icd.who.int/browse10/2019/en (дата обращения: 14.06.2024).
[2] Депрессивное расстройство (депрессия). URL: https://www.who.int/ru/news-room/fact-sheets/detail/depression (дата обращения: 21.04.2024).
[3] Катастрофа самолета Germanwings научила европейцев осторожности. URL: https://ria.ru/20150424/1060545533.html (дата обращения: 21.04.2024).

предпочтений в искусстве и культуре, а также полноте личных данных пользователя. Это позволит не только расширить знания о возможностях использования данных из социальных сетей для заблаговременной помощи лицам, страдающим депрессией, но и обеспечить безопасность окружающих, повысить эффективность предварительного скрининга сотрудников перед трудоустройством, а также реализовать механизм самодиагностики для лиц, заинтересованных в поддержании высокого уровня эмоционального здоровья [1].

Особенности сбора текстовых данных в русскоязычных социальных сетях

Прогнозирование депрессии на основе социальных сетей является активной областью исследований, где ученые стремятся выявить паттерны поведения и коммуникации пользователей, связанные с этим психологическим расстройством. Один из подходов включает в себя анализ текстовых данных, которые пользователи публикуют в своих профилях. В исследованиях показано, что определенные слова и фразы могут служить индикаторами депрессии. Например, использование слов, связанных с отрицательными эмоциями или состояниями, может быть признаком наличия депрессии у пользователя [2].

В ряде более старых работ, например, 2018 года, анализируются паттерны ввода текста на смартфонах для выявления признаков депрессии. Обнаружено, что образ жизни, выраженный через активности в социальных сетях и взаимодействия, может служить важным индикатором для определения психологического состояния пользователя [3].

В ряде работ прошлого десятилетия (например, 2017 года) активно используются зарубежные социальные сети из-за развитого API и простоты сбора данных, поскольку высокоуровневые API приложений из США и ЕС предоставляют простой механизм сбора данных без каких-либо ограничений, на языке программирования python разработано большое количество библиотек для успешного подключения и выгрузке всех необходимых данных [4]. В результате, большинство зарубежных исследований построено на исследовании семантического содержания публикаций пользователей социальной сети X (признана экстремистской организацией и запрещена на территории РФ) [5].

Важным аспектом является отличие типа вышеназванной социальной сети в сравнении с VK [6]. VK чаще используется в странах СНГ и обеспечивает широкий спектр функций, включая обмен сообщениями, публикацию фотографий и видеозаписей. X, напротив, ориентирован на создание коротких текстовых сообщений, известных как «твиты», которые ограничены 280 символами. Исходя из различий социальных сетей, можно отметить ограниченную релевантность зарубежных исследований по данной тематике.

В ряде исследований решалась задача прогнозирования депрессии, где с помощью библиотек в python посты в социальных сетях были проранжированы как позитивные и негативные с помощью оценки полярности, а точность прогнозирования тональности поста достигла 92,5 % [7].

Ряд исследователей в области прогнозирования интенсивности уровня депрессии на основе социальных сетей разделили все параметры на 5 основных групп, а далее тренировали нейросетевую модель LSTM для прогнозирования уровня депрессии [8], также они использовали аналогичную модель из двух групп факторов, собранных о пользователях с помощью зарубежных социальных сетей [9]. Авторы разделили все данные о пользователе на две группы: User-level features и Post-level features. В User-level features попали факторы, характеризующие пользователя с точки зрения его профиля:

количество сообщений, публичного взаимодействия с сообщениями других пользователей, а также наличие личных данных. Соответственно, в Post-level features входят данные непосредственно постов пользователей – их эмоциональная окраска, семантическое содержание, а для реализации модели была использована многослойная рекуррентная нейронная сеть. Модель, предложенная учеными, продемонстрировала точность в 89,5 % (метрика accuracy) в решении задачи бинарной классификации с условием сбалансированности классов, что говорит о возможности высокоэффективного прогнозирования уровня депрессии пользователя.

На данный момент существует несколько разработанных учеными моделей прогнозирования депрессии пользователей на основе данных социальных сетей: MDML (мультимодальная модель на основе использования словарей в исследовании 2017 года [10]), метод опорных векторов, сверточные нейронные сети, логистическая регрессия (и решение задачи классификации путем ввода отсечки), а также байесовский классификатор. Для имплементации модели была выбрана логистическая регрессия, поскольку она обеспечивает высокую точность, возможность работы с большим количеством параметров, а также управление порогом классификации (при решении задачи бинарной классификации, что соответствует цели исследования).

Сбор данных осуществлялся с помощью VK API, который имеет ряд ограничений. VK API не дает возможность сплошного сбора текстовых данных со страниц пользователей социальной сети ВКонтакте, а только с прямым указанием конкретного пользователя для сбора данных со страницы, поэтому обеспечить случайность выборки пользователей удалось через двухуровневый поиск «друзей». Так, сначала алгоритм собирал информацию о друзьях пользователя, а далее – о друзьях друзей (друзья «второго уровня»). Для обеспечения сбора данных VK требует создание приложения, поэтому было создано приложение, к которому был сгенерирован специальный токен доступа, позволивший обеспечить сбор данных.

Схема реализации сбора данных через друзей второго уровня указана на Рисунке 1.

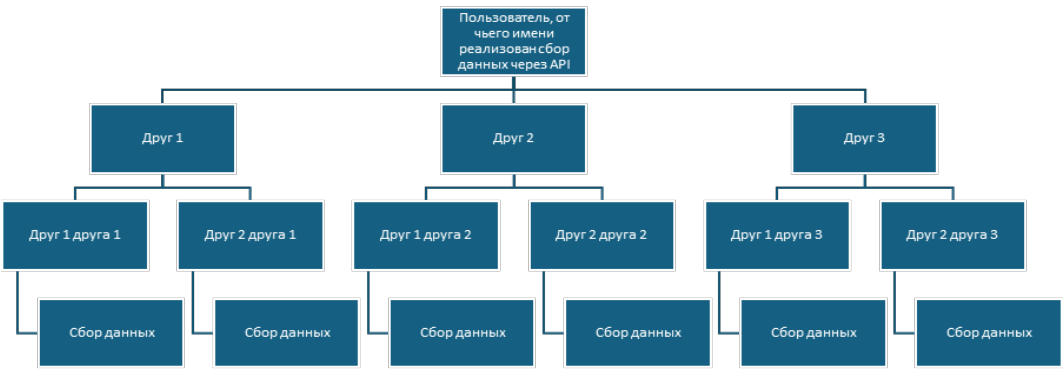


Рисунок 1 – Сбор данных через «друзей» второго уровня
Figure 1 – Data mining via 2nd level users

В результате имплементации алгоритма сбора данных удалось получить данные со страниц около 49,6 тысяч пользователей с общим объемом около 584 тысячи постов (около 11–12 постов на одного пользователя в среднем), но из-за ограничений API алгоритм имеет значительную временную сложность, его выполнение заняло более 10 часов. При наличии более мягких условий со стороны API возможно существенное расширение объема выборки. Тем не менее, действующие правила позволили собрать

значительную для исследования выборку постов и иных данных со страниц пользователей социальной сети.

На этапе предобработки данных все текстовые публикации каждого пользователя были объединены в один вектор и далее векторизованы методом TF-IDF. Аналогичным образом текстовые данные, указанные пользователями – статус, любимые фильмы, музыка, род деятельности – были объединены в отдельный вектор, состоящий из слов, добровольно указанных пользователями в описании «о себе» на личной странице в социальной сети. Пользователям был также добавлен признак, живут ли они в родном городе или указанный родной город отличается от указанного места жительства на основании сопоставления города, указанного пользователем как город проживания и как родной город.

После сбора и предобработки данных, вводными показателями для модели стали:

1. Наличие фотографии у пользователя (бинарный признак).
2. Статус пользователя (например, «online», «offline», «away» и т. д.).
3. Количество подписчиков у пользователя.
4. Пол пользователя (1 – мужской, 2 – женский).
5. Род занятий пользователя (например, студент, работник, пенсионер и т. д.).
6. Информация о интересах и предпочтениях пользователя.
7. Форма образования пользователя (например, очная, заочная и т. д.).
8. Признак того, что пользователь живет в родном городе (бинарный признак).
9. Количество постов, опубликованных пользователем.
10. Текстовая информация, например, тексты постов пользователя.
11. Статистика по текстовым данным (количество символов, слов, средняя длина поста).
12. Среднее значение количества указанных пользователем в профиле социальных сетей.

Для того, чтобы пометить пользователя признаком, является ли он депрессивным, использовался поиск на содержания сочетания «депресс» среди всех написанных им постов, при этом данные упоминания были исключены на этапе предобработки.

Следуя результатам работ [11, 12], будем считать, что факт упоминания депрессии у себя в социальных сетях свидетельствует о наличии депрессивного настроения с высокой вероятностью, поэтому любой пользователь, упоминавший депрессию у себя на странице, был помечен как депрессивный пользователь. В рамках задачи предобработки были устранены все упоминания депрессии, чтобы избежать переобучения модели.

С учетом вводных данных, формализация модели логистической регрессии представлена ниже.

$$P(y=1|x)=\frac{1}{1+e^{-(\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_6x_6+\beta_7x_7+\beta_8x_8+\beta_9x_9+\beta_{10}x_{10}+\beta_{11}x_{11}+\beta_{12}x_{12})}}. (1)$$

В рамках формулы x – это вышеперечисленные 12 вводных факторов с соответствующими им весовыми коэффициентами.

После предобработки и выделения всех пользователей было выявлено, что признаки депрессии были выявлены у около 550 пользователей, остальные пользователи – 49246 человек – были идентифицированы как лица без признаков депрессии.

Учитывая серьезную проблему дисбаланса классов (всего 0,1 % всех пользователей относился к классу пользователей с депрессией), была проведена ребалансировка путем сокращения количества пользователей недепрессивных пользователей случайным образом в 1000 раз, до 500 пользователей.

Логистическая регрессия используется для решения задачи бинарной классификации, в которой предсказывается вероятность принадлежности объекта к одному из двух классов. В контексте решения задачи классификации по прогнозированию депрессии это «депрессивные пользователи» и «недепрессивные пользователи». Модель логистической регрессии предсказывает вероятность, что объект X принадлежит классу 1 (в условиях исследования – имеет ли пользователь признаки депрессии). Также были использованы модели градиентного спуска и случайного леса и подобрана оптимальная граница разделения на классы.

Применение моделей машинного обучения для прогнозирования депрессии

По результатам работы, выборка была разделена в пропорции 70/30 как тренировочная и тестовая – для проверки надежности и качества модели. При имплементации алгоритма на 1000 пользователях были получены высокие показатели качества решения задачи классификации.

Были реализованы следующие модели: логистическая регрессия, градиентный спуск, случайный лес.

Для моделей градиентного спуска был осуществлен подбор гиперпараметров с помощью языка программирования Python и модуля GridSearchCV из библиотеки scikit-learn. Наилучшая модель в рамках GradientBoostingRegressor продемонстрировала лучшую точность метрики roc_auc (0.98) при шаге обучения в 0.1 и количестве оценочных параметров в 50. Для случайного леса наилучшая точность была при максимальной глубине дерева в 10 и количестве оценочных параметров в 200.

Для выбора границы разделения на классы были построены графики зависимости границы разделения на классы и оценки ROC-AUC и представлены на Рисунке 2. На Рисунке 3 представлены результаты бинарной классификации.

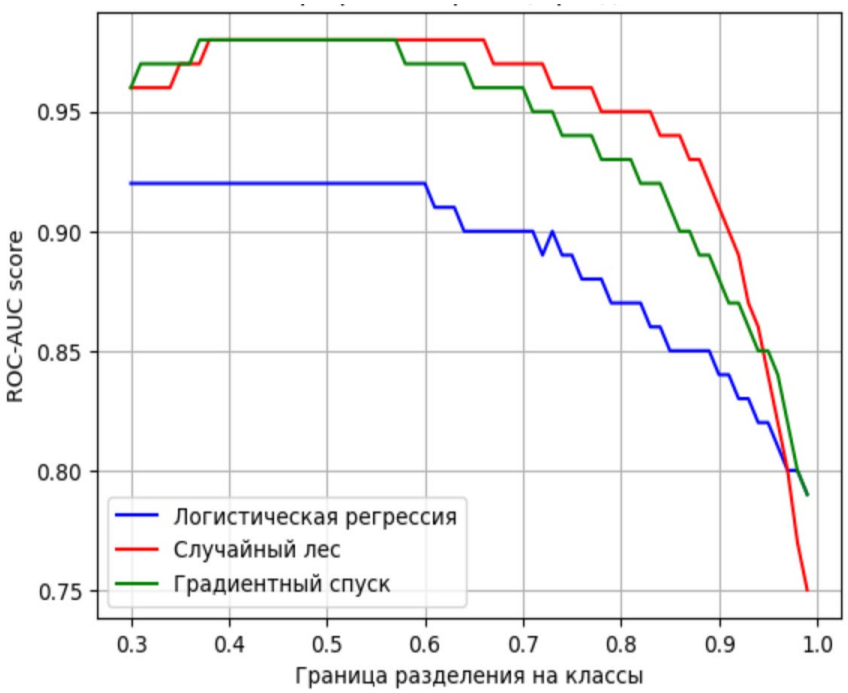


Рисунок 2 – Оценка ROC-AUC при разной границе разделения на классы
Figure 2 – ROC-AUC score compared with different classification threshold

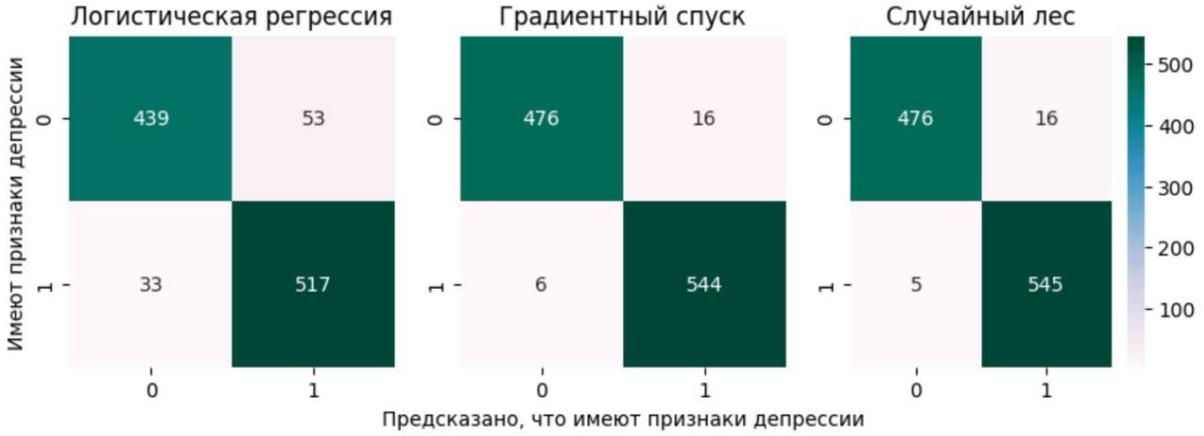


Рисунок 3 – Матрица результатов по классификации пользователей как имеющих признаки депрессии
Figure 3 – Result matrix for classifying users as having signs of depression

Согласно Рисунку 2, оптимальной границей разделения на классы является 0,5, поскольку в таком случае для всех трех моделей метрика ROC-AUC максимальна.

Для класса пользователей без признаков депрессии точность (precision) составляет 0,92. Это означает, что из всех пользователей, которых модель предсказала как без признаков депрессии, 92 % действительно не страдают от депрессии. Для класса пользователей с депрессией точность составила 0,91. Это означает, что из всех пользователей, которых модель предсказала как депрессивные, 91 % действительно являются депрессивными.

Для класса пользователей без депрессии полнота (recall) составляет 0,90. Это означает, что моделью правильно предсказано из всех пользователей 90 %, не страдающих от депрессии. Для противоположного класса полнота равна 0,94. Это означает, что моделью правильно определено 94 % депрессивных пользователей из всех пользователей.

Мера F1 превысила 0,9, что говорит о высокой предсказательной силе модели. Общая точность модели (accuracy) равна 0,92. Это означает, что моделью были правильно классифицированы 92 % всех пользователей.

В рамках исследования были также использованы два других метода – RandomForestRegressor и GradientBoostingRegressor, в результате точность для первого метода составила 0,98, для второго аналогично 0,98, что говорит о преимуществах в использовании ансамблевых моделей, результаты представлены в Таблице 1.

Таблица 1 – Сравнение эффективности моделей классификации
Table 1 – Comparison of efficiency of the classification models

Модель	Accuracy	Precision	F1	ROC-AUC
Логистическая регрессия	0,92	0,93	0,92	0,92
RandomForestRegressor	0.98	0,99	0,98	0,98
GradientBoostingRegressor	0,98	0,99	0,98	0,98

На основе матрицы результатов можно сделать выводы о высоком качестве моделей градиентного спуска и случайного леса, а также низкой ошибке второго рода, что делает решение задачи эффективным и точным.

Также сопоставление моделей можно осуществить с помощью модели ROC-AUC (Рисунок 4), которая демонстрирует примерно одинаково высокую точность градиентного спуска и случайного леса.

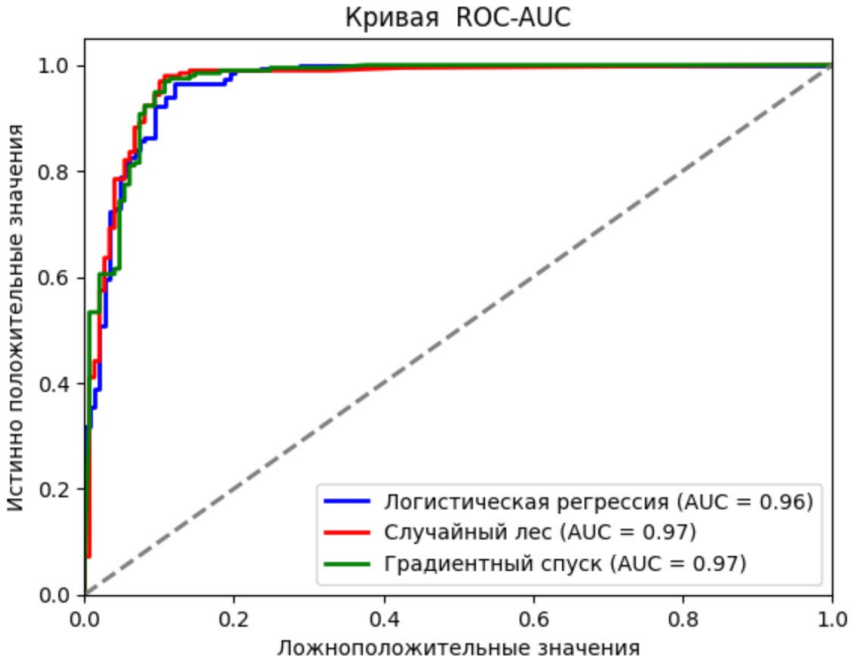


Рисунок 4 – Кривая ROC-AUC для моделей прогнозирования признаков депрессии
Figure 4 – ROC-AUC curve for forecasting of depressed users

Таким образом, прогнозирование депрессии на основе пользовательской информации в русскоязычной социальной сети возможно с высокой точностью даже с использованием простой модели логистической регрессии для классификации, так и использованием продвинутых моделей машинного обучения – случайного леса и градиентного спуска – для более высокого качества моделей.

Сравнительная характеристика депрессивных и недепрессивных пользователей

Несмотря на достигнутые результаты, модель требует доработки и дальнейшего улучшения. Важной перспективой исследования является включение других методов машинного обучения, таких как нейронные сети, для повышения точности и обобщающей способности модели. Также важно учитывать этические аспекты использования данных пользователей и обеспечивать их конфиденциальность и защиту.

Модель показывает высокую точность, не сильно уступающую современным зарубежным нейросетевым моделям, которые, в свою очередь, проигрывают логистической регрессии в сложности управления, высокой трудоемкости и сложности архитектуры. На данный момент зарубежные модели прогнозирования депрессии на основе англоязычных постов демонстрируют точность (accuracy) в районе 95 % [5].

Сравнение депрессивных и недепрессивных пользователей по показателям числа постов (n_posts), средней длины поста (post_len), количества слов в посте (word_count) и числа подписчиков (followers_count) позволяет выявить особенности в их поведении и характеристиках их активности в VK.

Депрессивные пользователи имеют в среднем значительно большее количество постов (77) по сравнению с пользователями без выявленных признаков депрессии (11).

Это может указывать на более высокую активность в социальной сети у пользователей, страдающих депрессией.

Пользователи, страдающие от депрессии, используют больше слов в своих постах (в среднем 8558 слов за последние 100 постов) по сравнению с пользователями без спрогнозированной депрессии (340 слов). Это указывает на более детальное и развернутое выражение своих мыслей и эмоций в сообщениях [4, 13].

Сравнение пользователей, исходя из выявления у них признаков депрессии, приведено в Таблице 2.

Таблица 2 – Сравнение пользователей с выявленными признаками депрессии и без них
Table 2 – Comparison of users with and without identified signs of depression

Есть признаки депрессии	Среднее количество публикаций	Общее количество символов	Общее количество слов	Средняя длина публикации
Нет	11	2306	340	16
Да	77	56716	8559	104

Регрессия представляет собой метод не только для прогнозирования непрерывной величины, но и для оценки вероятности события. В контексте прогнозирования депрессии у пользователей социальной сети, в частности, логистическая регрессия может быть использована для оценки вероятности наличия у пользователя депрессии на основе его данных из профиля и текстовых постов на личных страницах.

Оценка вероятности депрессии у пользователя и ее выявление может быть полезной в практических целях. Например, на основе такой оценки можно выделить пользователей с высокой вероятностью депрессии для более тщательного мониторинга и предоставления медицинской и психологический помощи.

Одним из преимуществ регрессии является возможность оценки вероятности не только на момент времени, но и в динамике. Это означает наличие возможности анализировать изменения вероятности депрессии у пользователя со временем на основе его новых данных из социальной сети. Такой подход позволяет более точно отслеживать и прогнозировать психологическое состояние пользователей и своевременно предпринимать необходимые меры для поддержки и помощи.

В качестве дальнейшего использования модели логистической регрессии возможно выделить следующие аспекты:

1. Использование данных об аватаре пользователя. Аватар может содержать ценную информацию о настроении, эмоциональном состоянии и поведенческих паттернах пользователя. Автоматический анализ аватаров с использованием компьютерного зрения и алгоритмов машинного обучения может помочь выявить характеристики, которые связаны с депрессией, такие как выражение лица, поза и цветовая гамма.

2. Вместе с традиционными методами машинного обучения внедрение нейросетевых методов прогнозирования поможет повысить точность модели, поскольку нейронные сети могут выявлять сложные нелинейные зависимости между признаками и целевой переменной. Это может привести к более точным и эффективным моделям прогнозирования, способным учитывать разнообразные и сложные характеристики пользователей.

3. Для улучшения качества прогнозирования в будущем возможно рассмотрение подхода, основанного на прогнозировании депрессии на различных группах факторов с последующим объединением результатов в единую метрику. Например, путем

разделения факторов на социальные, психологические, биологические и дальнейшего агрегирования полученных данных.

4. Исходя из проблем задачи прогнозирования депрессии, в перспективе важно учитывать изменения в психическом состоянии пользователя со временем и анализировать их динамику. Использование алгоритмов регрессии, таких как временные ряды или рекуррентные нейронные сети, может помочь моделировать эволюцию депрессии у пользователя и предсказывать вероятные тенденции в его психическом здоровье.

Заключение

Прогнозирование и выявление депрессии пользователей на основе семантических данных из постов в социальной сети ВКонтакте и данных их профиля с использованием логистической регрессии представляет собой эффективный и актуальный метод анализа психологического состояния пользователей в онлайн-среде.

Описанный в работе подход позволяет автоматизировать и улучшить процессы выявления и поддержки пользователей, страдающих депрессией. Предсказание депрессии на основе данных социальных сетей может помочь в разработке персонализированных интервенций и предложений поддержки для пользователей, а также в повышении эффективности программ по борьбе с депрессией в социальных сетях.

Потенциальными пользователями разработанных в исследовании средств могут быть кадровые подразделения коммерческих организаций, где психологическое здоровье сотрудников играет особо важную роль (например, в авиации у пилотов). Такое решение позволит осуществлять постоянный мониторинг состояния сотрудников, а также проверять психологическое состояние при приеме на работу.

Реализация регрессионных алгоритмов при анализе семантического содержания публикаций может быть полезна для исследования влияния социальной сети на психическое здоровье пользователей. Анализ текстовых данных из постов пользователей и их профилей может помочь выявить паттерны и факторы, способствующие развитию депрессии, что, в свою очередь, может быть использовано для разработки профилактических мер и улучшения условий онлайн-взаимодействия.

Таким образом, развитие исследования выявления и прогнозирования депрессии на основе данных социальных сетей представляет собой перспективную и важную область, которая может принести значительный вклад в области психологии, здравоохранения и информационных технологий. Реализация предложенных перспектив позволит повысить эффективность и точность модели, что окажет влияние на развитие методов прогнозирования депрессивного расстройства личности.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Тарабакина Л.В. *Эмоциональное здоровье подростка: риски и возможности.* Москва: Московский педагогический государственный университет; 2017. 194 с. Tarabakina L.V. *Emotsional'noe zdorov'e podrostka: riski i vozmozhnosti.* Moscow: Moscow Pedagogical State University; 2017. 194 p. (In Russ.).
2. Richter T., Richter-Levin G., Okon-Singer H., Fishbain B. Machine Learning-Based Behavioral Diagnostic Tools for Depression: Advances, Challenges, and Future Directions. *Journal of Personalized Medicine.* 2021;11(10). https://doi.org/10.3390/jpm11100957

3. Rajkomar A., Oren E., Chen K. et al. Scalable and accurate deep learning with electronic health records. *npj Digital Medicine.* 2018;1(1). https://doi.org/10.1038/s41746-018-0029-1
4. Браницкий А.А., Шарма Я.Д., Котенко И.В., Федорченко Е.В., Красов А.В., Ушаков И.А. Определение психического состояния пользователей социальной сети Reddit на основе методов машинного обучения. *Информационно-управляющие системы.* 2022;(1):8–18. https://doi.org/10.31799/1684-8853-2022-1-8-18
Branitskiy A.A., Sharma Y.D., Kotenko I.V., Fedorchenko E.V., Krasov A.V., Ushakov I.A. Determination of the mental state of users of the social network Reddit based on machine learning methods. *Informatsionno-upravlyayushchie sistemy = Information and Control Systems.* 2022;(1):8–18. (In Russ.). https://doi.org/10.31799/1684-8853-2022-1-8-18
5. Suhara Y., Xu Y., Pentland A.S. DeepMood: Forecasting Depressed Mood Based on Self-Reported Histories via Recurrent Neural Networks. In: *WWW '17: Proceedings of the 26th International Conference on World Wide Web, 3-7 April 2017, Perth, Australia.* Switzerland: International World Wide Web Conferences Steering Committee; 2017. P. 715–724. https://doi.org/10.1145/3038912.3052676
6. Каплун И.Г., Герасимов П.Е., Чучин В.В., Клюшников Н.В. Особенности коммуникативной компетентности студентов – активных пользователей социальных сетей. *Скиф. Вопросы студенческой науки.* 2022;(11):163–167.
Kaplun I.G., Gerasimov P.E., Chuchin V.V., Klyushnikov N.V. Features of the communicative competence of students – active users of social networks. *Skif. Voprosy studencheskoi nauki = Sciff. Questions of Students Science.* 2022;(11):163–167. (In Russ.).
7. Govindasamy K.A., Palanichamy N. Depression Detection Using Machine Learning Techniques on Twitter Data. In: *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 6-8 May 2021, Madurai, India.* IEEE; 2021. P. 960–966. https://doi.org/10.1109/ICICCS51141.2021.9432203
8. Ghosh S., Anwar T. Depression Intensity Estimation via Social Media: A Deep Learning Approach. *IEEE Transactions on Computational Social Systems.* 2021;8(6):1465–1474. https://doi.org/10.1109/TCSS.2021.3084154
9. Zogan H., Razzak I., Wang X. et al. Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media. *World Wide Web.* 2022;25(1):281–304. https://doi.org/10.1007/s11280-021-00992-2
10. Shen G., Jia J., Nie L., Feng F., Zhang C., Hu T., Chua T.-S., Zhu W. Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution. In: *IJCAI'17: Proceedings of the 26th International Joint Conference on Artificial Intelligence, 19-25 August 2017, Melbourne, Australia.* AAAI Press; 2017. P. 3838–3844. https://doi.org/10.24963/ijcai.2017/536
11. Углова А.Б., Низомутдинов Б.А. Анализ деструктивного контента телеграмм-каналов как фактора развития саморазрушающего поведения. *International Journal of Open Information Technologies.* 2022;10(11):81–86.
Uglova A.B., Nizomutdinov B.A. Analysis of the destructive content of telegram channels as a factor in the development of self-destructive behavior. *International Journal of Open Information Technologies.* 2022;10(11):81–86. (In Russ.).
12. Ковпак Д.В. Когнитивно-поведенческая терапия суицидального поведения. *Вестник Московской международной академии.* 2021;(2):55–63.
Kovpak D.V. Cognitive behavioral therapy of suicidal behavior. *Vestnik Moskovskoi mezhdunarodnoi akademii = Bulletin of the Moscow International Academy.* 2021;(2):55–63. (In Russ.).

13. Латынов В.В., Овсянникова В.В. Прогнозирование психологических характеристик человека на основании его цифровых следов. *Психология. Журнал высшей школы экономики.* 2020;17(1):166–180. https://doi.org/10.17323/1813-8918-2020-1-166-180
Latynov V.V., Ovsyannikova V.V. Predicting Psychological Characteristics from Digital Footprints. *Psikhologiya. Zhurnal vysshei shkoly ekonomiki = Psychology. Journal of Higher School of Economics.* 2020;17(1):166–180. (In Russ.). https://doi.org/10.17323/1813-8918-2020-1-166-180

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Солохов Тимур Дамирович, магистр, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация.
e-mail: TDSolokhov@fa.ru
ORCID: 0009-0001-6073-3754

Timur D. Solokhov, master, Financial University under the Government of the Russian Federation, Moscow, the Russian Federation.

Кочкаров Азрет Ахматович, доктор технических наук, доцент, профессор кафедры анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация.
e-mail: AAKochkarov@fa.ru
ORCID: 0000-0002-3232-5331

Azret A. Kochkarov, Doctor of Technical Sciences, Associate Professor, Professor of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, Moscow, the Russian Federation.

Статья поступила в редакцию 31.05.2024; одобрена после рецензирования 13.06.2024; принята к публикации 24.06.2024.

The article was submitted 31.05.2024; approved after reviewing 13.06.2024; accepted for publication 24.06.2024.