

УДК 004.9, 004.5, 004.41/.42, 004.43
МРНТИ 50.41.25

РАННЕЕ ВЫЯВЛЕНИЕ СКЛОННОСТИ ДЕПРЕССИВНОГО ПОВЕДЕНИЯ У ПОДРОСТКОВ ЧЕРЕЗ СОЦИАЛЬНЫЕ СЕТИ

А. ДЖАНТЫКОВ^{1,2}, Х. НУРЛЫБЕКОВ^{1,2}, К. МУРАТОВА^{1,2}, Б. ОМАРОВ^{1,2}

¹Казахстанская инновационная лаборатория при поддержке фонда ЮНИСЕФ

²Международный университет информационных технологий

Аннотация: Благодаря анонимности интернет-СМИ и социальных сетей, люди склонны выражать свои чувства и страдания в интернет-сообществах. Для предотвращения нежелательных эффектов со стороны подростков необходимо выявлять сообщения о депрессивных поведении и представления пользователей о самоубийствах в киберпространстве методами обработки естественного языка. Мы фокусируемся на социальной сети «ВКонтакте» и классифицируем сообщения пользователей с потенциальным самоубийством и без суицидального риска с помощью обработки текстов и методов, основанных на машинном обучении.

Ключевые слова: депрессивное поведение, социальные сети, анализ данных

EARLY DETECTION OF DEPRESSIVE BEHAVIOR IN ADOLESCENTS THROUGH SOCIAL NETWORKS

Abstract. Due to the anonymity of the Internet and social networking sites, people are expressing their senses and in the Internet community. For those who do not want to be self-reported, they have to report on cyberbullying in cyberspace and to represent users in cyberspace methods. We focus on the social networking site “VKontakte” and classify users with potential suicidal and suicidal behaviors, with the help of text processing and techniques, based on machine learning.

Keywords: Depressive behavior, Social networks, Data analysis

ӘЛЕУМЕТТІК ЖЕЛІЛЕР АРҚЫЛЫ ЖАСӨСПІМДЕРДЕГІ ДЕПРЕССИВТІ МІНЕЗ-ҚҰЛЫҚТЫ АНЫҚТАУ

Аңдатпа: Ғаламтор-БАҚ және әлеуметтік желілердің құпиялығына байланысты, интернет-қоғамдастықтардағы адамдар өз сезімдерін және жағдайларын интернет желілерінде білдіруге тырысады. Өзін-өзі құрбандыққа шалудың алдын алу үшін, алдымен сол сарындағы хабарларды анықтау керек. Біз үлкен деректер жинауды және талдауды, сонымен қатар, машинаны үйрету әдістерін қолдана отырып, «ВКонтакте» әлеуметтік желісінде, жан күйзелісіне түскен депрессивті хабарламалар иелерін анықтауға тырыстық.

Түйінді сөздер: депрессия, күйзеліс, әлеуметтік желілер, деректерді талдау

ВВЕДЕНИЕ

Развитие Интернета и современных средств электронной коммуникации, доступность и распространенность социальных сетей, простота использования социальных платформ и обеспечиваемые ими возможнос-

ти для самовыражения делают социальные сети чрезвычайно привлекательными для современной молодежи. Растущая популярность и возрастающее многообразие площадок для коммуникации в Интернет имеют как

позитивные, так и негативные последствия для молодых людей и всего общества в целом.

Современное общество характеризует- ся кризисными явлениями во многих сферах жизни. Трудные социально-экономические условия в значительной мере осложняют реальность. Происходят резкая смена привычных стереотипов, дестабилизация положения в обществе и окружающей среде. Эти и многие другие факторы определяют нарушения, связанные с особой ранимостью современных подростков, возникает страх перед действительностью с последующей социальной дезадаптацией. При отсутствии своевременной психологической помощи такие дети могут совершить суицидальную попытку. Поэтому углубленное изучение факторов, обуславливающих депрессивное поведение, особенностей девиантного поведения подростков и средств коррекции суицидального поведения является необходимой составляющей для создания эффективной системы профилактики подростковых суицидов.

ОБЗОР ЛИТЕРАТУР

Из ежегодно совершаемых в мире 10-20 млн. попыток самоубийств около 50% приходится на пациентов, страдающих депрессией.

Насколько распространены депрессивные расстройства?

Депрессивными расстройствами страдают миллионы жителей нашей планеты. Распространенность депрессивных расстройств в популяции к 90-м годам XX века в развитых странах Европы и в США достигла 5-10%. На

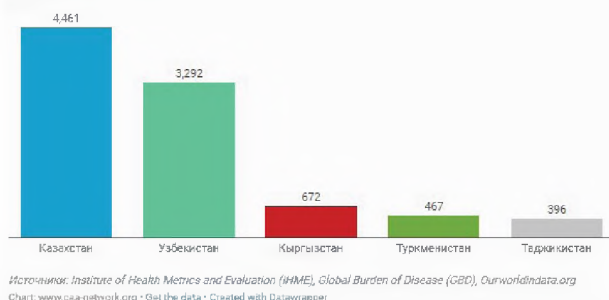
сегодняшний день самое крупное международное эпидемиологическое исследование по изучению частоты психических расстройств в общемедицинской практике было проведено Всемирной организацией здравоохранения (ВОЗ) в конце 1980-х годов. Оно называлось Psychological Disorders in Primary Care и проводилось в 18 странах Америки, Европы и Азии. Согласно данным этого исследования различные расстройства психической сферы отмечаются у каждого четвертого (24%) пациента общей сети здравоохранения, расстройства депрессивного спектра присутствуют у каждого пятого (21%) пациента. Из расстройств депрессивного спектра чаще всего отмечаются собственно депрессивные (10,4%) и тревожные (10%) расстройства. Показатели распространенности психоэмоциональных расстройств значительно варьируют в зависимости от страны и колеблются от 20% в Китае до 50% в Чили.

Согласно мировым статистическим данным, Казахстан занимает 3-4 место в мире по количеству суицида [1], а также 1 место по количеству самоубийств среди девушек в возрасте от 15 до 19 лет во всем мире [2][6].

Данные графики в Рисунке 1 подтверждают высокий уровень суицида в стране по сравнению с соседними Азиатскими странами:

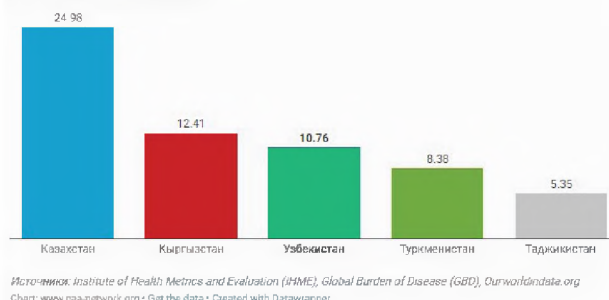
Суицидальные мысли - это мысли людей о самоубийстве. Его можно рассматривать как показатель риска суицида. Суицидальные мысли включают мимолетные мысли, обширные мысли, подробное планирование,

Число самоубийств
(2016 г.)



а) Число самоубийств
в центрально-азиатских странах в 2016 году

Уровень самоубийств
(на 100 000 населения, 2016 г.)



б) Уровень самоубийств на 100 тыс. населения в
центрально-азиатских странах в 2016 году

Рис. 1 – Уровень самоубийств в центрально-азиатских странах в 2016 году

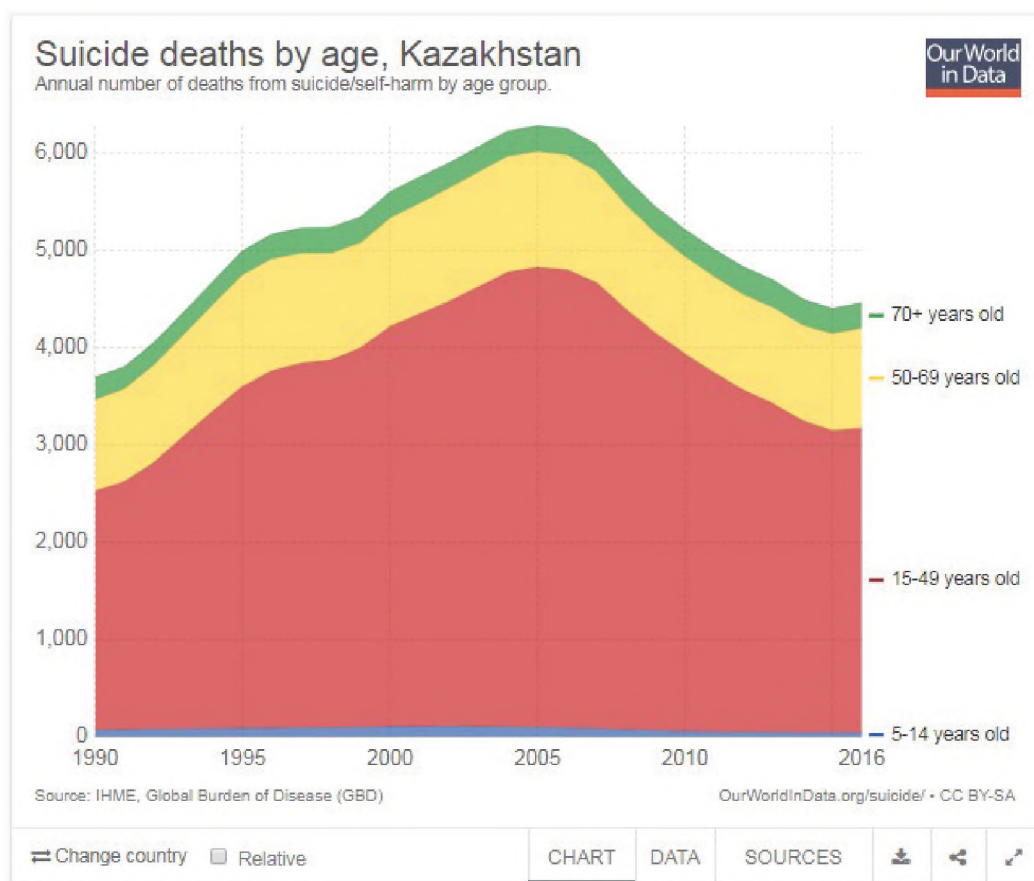


Рис. 2 – Уровень самоубийств по возрасту

ролевые игры и неполные попытки. Согласно отчету, ВОЗ [3], в 2015 г. по оценкам 788 000 человек во всем мире совершили самоубийство. И большое количество людей, особенно подростков, были зарегистрированы с суицидальными идеями. Таким образом, одним из возможных подходов к эффективному предотвращению самоубийств является раннее выявление суицидальных идей.

МЕТОДЫ

Ключевые слова, подтверждающие возможность определения поста как суицидального

Что значит «ключевые слова, подтверждающие возможность определения поста как суицидального»? Есть некий набор слов, который зачастую используют люди, которые вздумали совершить суицид. В основном, эти слова напрямую связаны с идеей жизни и смерти, но порой бывает, что в посты, которые пишутся людьми, находящимися в поникшем, суицидальном настроении,

стараясь избежать использования слов, напрямую означающих их попытку суицида. Но они стараются использовать синонимы этим самым словам, тем самым давая нам возможность отыскивать их посты, используя все более новые и новые наборы ключевых слов.

Из предыдущей темы были выявлены ключевые слова, ассоциирующиеся с самоубийством. Например, гроб, жизнь, конец и т.д. Данные ключевые слова помогут в поиске суицидальных постов в социальных сетях.

По мере нахождения суицидальных постов, база ключевых слов будет пополняться, тем самым обеспечивая более точно определение суицидальных постов.

Для более точного нахождения суицидальных постов, по ключевым словам, будет произведена встреча с экспертами в этой области, для расширения базы ключевых слов и понимания характерной подписи самоубийц.

Набор имеющихся ключевых слов: убейте, умереть, прощайте, жизнь говно, гроб, не

люблю себя, ненавижу себя, хочу умереть, повешаться, суицид, самоубийство, неразделенная любовь, ненавижу школу, умрите, сдохните, сдохну, не хочу жить, на тот свет, умру, рай, ад, я виноват, депрессия, умираю и др.

Разработка TelegramBot, для сбора постов

Для сбора данных в социальной сети был разработан TelegramBot, который собирает посты, публикуемые в режиме реального времени, по ключевым словам, найденным из предыдущей темы.

Данный бот собирает все найденные им данные в специальный файл, в котором хранится текст поста, ссылка на пост, ссылка на автора поста.

Бот работает круглосуточно, что позволяет находить посты независимо от часового пояса того, кто захочет опубликовать пост, в котором находится ключевое слово, которое было введено.

Бот использует сразу две технологии одновременно, это TelegramAPI и APIVkontakte. На основе StreamingAPI (API для аналитики от «ВКонтакте»), боту направляется 1% публикуемой информации (смотреть документацию к StreamingAPI [5]) в социальной сети «ВКонтакте», в момент публикации бот отправляет информацию разработчику, попутно сохраняя ее.

Разработка алгоритма машинного обучения на основе собранных данных

По итогу сбора около 5 тыс. постов с возможной склонностью к суициду, будет проведена ручная очистка данных, а также классификация на суицидальные и несуицидальные посты.

Вскоре после этого будет разработан алгоритм машинного обучения, который на основе очищенных и классифицированных на суицидальные посты данных будет автоматически классифицировать новые найденные посты на возможность к определению поста как носящего суицидальный характер, а также на основе суицидальных постов, алгоритм будет пополнять базу ключевых слов,

для более точного нахождения суицидальных постов, выдаваемых ботом.

Анализ пользователей

После определения суицидального поста, начинается анализ пользователя, автора либо участника соответствующей группы на склонность к суициду. Проводится анализ предыдущих опубликованных записей на странице пользователя, проводится анализ на подписку на группы, которые имеют суицидальные наклонности, проводится анализ пользователя на различные факторы возможности суицида.

После определения пользователя как возможного самоубийцу, информация о нем будет передана в учебное заведение, в котором он обучается, для последующей работы с ним и для предотвращения суицида. Но прежде чем реализовать эту возможность, требуется запросить данные обо всех учащихся на территории Казахстана. После определения потенциального самоубийцы, будет произведен поиск по имеющийся базы данных всех учеников на соответствие имени и фамилии. Но так как многие дети предпочитают вести, так называемые, «фейковые» аккаунты, могут возникнуть некоторые трудности с точным определением местонахождения потенциального самоубийцы. Для этого был предусмотрен вариант написания данному пользователю через социальную сеть «ВКонтакте», а также попытка переубедить его совершать суицид в режиме «онлайн». Данная попытка может оказаться малоэффективной, так как мало кто захочет слушать незнакомого человека, особенно когда тот желает поменять свое мнение насчет того или иного вопроса. Но долгое изучение самоубийц говорит о том, что внимание со стороны других людей к их проблеме способно их переубедить и начать жизнь по-новому.

Если же пользователь оставил свои настоящие контактные данные и мы в состоянии определить его местонахождение через базу данных учебных заведений, то информация о пользователе будет передана напрямую в то учебное заведение, в котором обучается по-

тенциальный самоубийца. Далее, с ним будет работать психолог учебного заведения либо классный руководитель (куратор группы).

Это поможет преждевременно предотвратить попытку суицида.

РЕЗУЛЬТАТЫ

Данные, которые сохраняет бот:

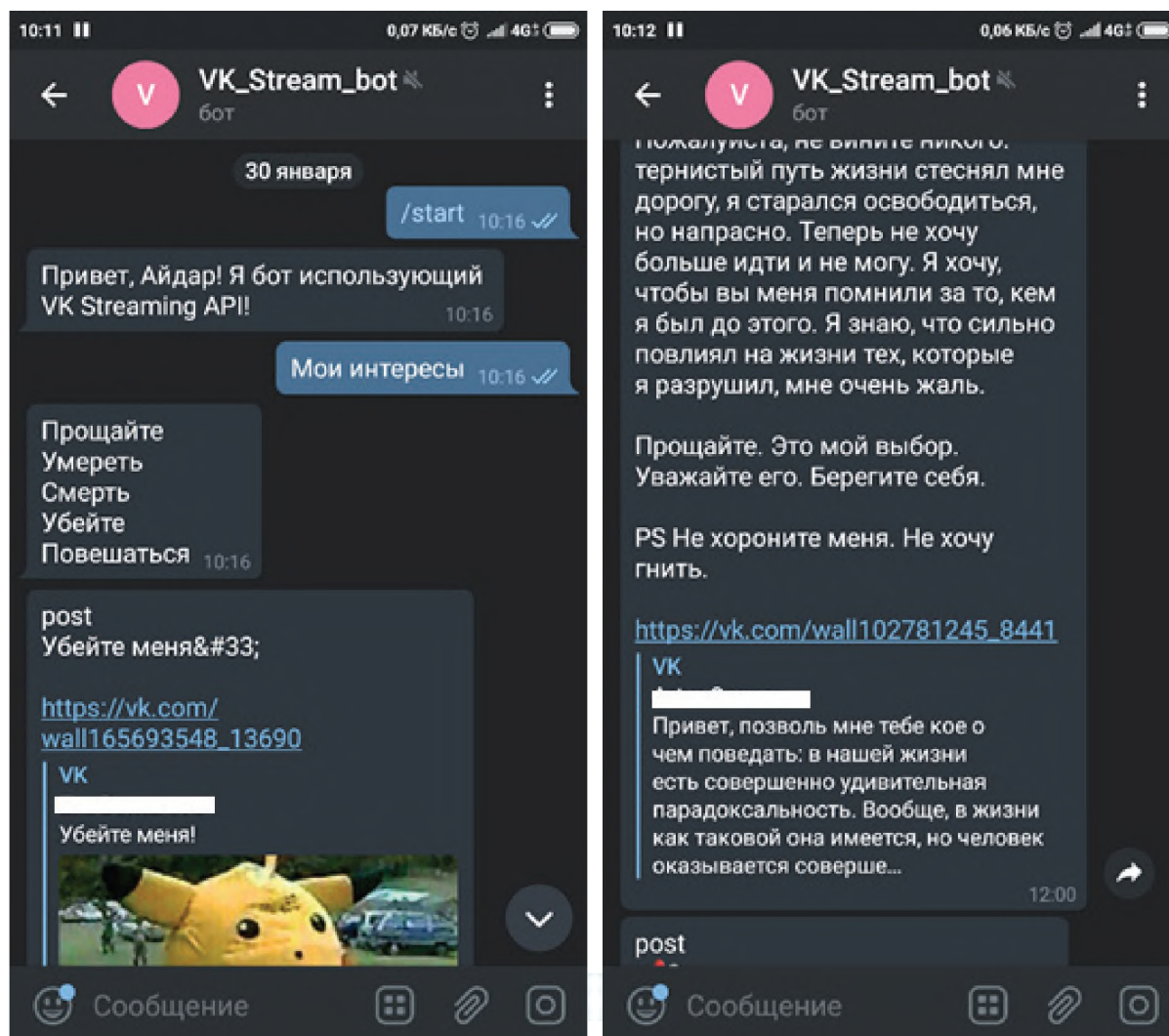


Рис. 3 – Пример работы телеграм-бота

То, как выглядит файл, в котором сохраняются все данные, собранные ботом:

[illegible]

Рис. 4 – Пример сбора данных

То, как будут выглядеть данные, после ручной очистки и классификации:

	A	B
1	Not suitable posts	Suitable posts
2	https://vk.com/wall341280824_2903	https://vk.com/wall102781245_8441
3	https://vk.com/wall326478360_4590	https://vk.com/wall165693548_13690
4	https://vk.com/wall326672312_5207	
5	https://vk.com/wall-175647646_132	
6	https://vk.com/wall61319882_3477	
7	https://vk.com/wall-110065990_3826	
8	https://vk.com/wall33437158_2973	
9	https://vk.com/wall-149817246_979	
10	https://vk.com/wall136638779_18402	
11	https://vk.com/wall438176878_124	
12	https://vk.com/wall316396032_2615	
13	https://vk.com/wall141870471_8564	

Рис. 5 – Классификация данных

Пример работы бота:

	Айдар /start	1:00:33 PM
	VK_Stream_bot Привет, Айдар! Я бот использующий VK Streaming API!	1:00:34 PM
	Айдар Добавить	1:00:41 PM
	VK_Stream_bot Что добавить?	1:00:42 PM
	Айдар смерть	1:00:47 PM
	VK_Stream_bot Successful	1:00:50 PM
	Айдар Добавить	1:00:52 PM
	VK_Stream_bot Что добавить?	1:00:52 PM
	Айдар хочу умереть	1:00:59 PM
	VK_Stream_bot Successful	1:01:01 PM
	Айдар Добавить	1:01:03 PM
	VK_Stream_bot Что добавить?	1:01:03 PM
	Айдар жизнь говно	1:01:09 PM

Рис. 6 – Пример работы бота

Данные отправляемые ботом пользователю (все данные сохраняются в отдельном файле):

	Айдар жизнь говно	1:01:09 PM
	VK_Stream_bot Successful	1:01:11 PM
	post Сложно сказать, что в жизни главное. Жизнь — она бессмысленна сама по себе. Просто надо найти себе дело, чтобы занять этот промежуток между рождением и смертью. Какой смысл жизни? Да нет его. Родить детей — это репродуктивная функция, она вне нас лежит. К чему стремиться? Количественные какие-то факторы человека вряд ли мотивируют. Два завтрака не съешь. Надо найти себе занятие. Просто ты придумал себе какую-то игру, и ты в нее играешь. https://vk.com/wall206086333_8162 VK Сложно сказать, что в жизни главное. Жизнь — она бессмысленна сама по себе. Просто надо найти себе дело, чтобы занять этот промежуток между рождением...	1:01:48 PM
	post Смерть пасажир: на трасі «Суми-Полтава» ДТП Смертельна аварія сталась на трасі «Суми-Полтава» сьогодні о дев'ятій годині ранку. Водій ВАЗу, обганяючи вантажівку, вискочив на зустрічну смугу та зіткнувся з автомобілем Volkswagen. Внаслідок цього іномарка наїхала на відбійник. Загинув пасажир. #Суми #Сумы #Summy	1:02:37 PM
	https://vk.com/wall-87696376_23482	

Рис. 7 – Пример нахождения потенциально-депрессивных постов

Данный способ сбора данных является очень эффективным, так как для сбора не требуется постоянный контроль над ситуацией, все происходит автоматически.

ЗАКЛЮЧЕНИЕ

Количество текста продолжает расти с популяризацией социальных сетей. И профилактика самоубийств остается важной задачей в нашем современном обществе. Поэтому необходимо разработать новые методы выявления в интернете текстов, содержащих суи-

цидальные идеи, в надежде на то, что суицид можно предотвратить.

В данной работе мы исследовали проблему обнаружения суицидальности в онлайн-контенте, генерируемом пользователями. Использование обработки больших данных из социальных сетей с помощью алгоритмов машинного обучения поможет существенно сократить попытки суицида среди подростков путем преждевременного определения суицидального настроения через социальные сети.

ЛИТЕРАТУРА

1. https://online.zakon.kz/Document/?doc_id=30468239
2. <https://www.dropi.ru/posts/15-stran-s-samym-vysokim-urovнем-samoubijstv-v-mire>
3. http://www.who.int/gho/mental_health/suicide_rates/en/
4. <https://batenka.ru/resource/suicide/message/>
5. https://vk.com/dev/streaming_api
6. <https://kursiv.kz/news/obschestvo/2018-02/kazakhstan-lidiruet-po-chislu-samoubiystv-v-central-noy-azii>