ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ
АКАДЕМИЯСЫ» РҚБ

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ

ХАБАРЛАРЫ

ИЗВЕСТИЯ

РОО «НАЦИОНАЛЬНОЙ
АКАДЕМИИ НАУК РЕСПУБЛИКИ
КАЗАХСТАН»

N E W S

OF THE NATIONAL ACADEMY
OF SCIENCES OF THE REPUBLIC
OF KAZAKHSTAN

SERIES OF PHYSICS AND MATHEMATICS

1 (353)

JANUARY – MARCH 2025

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

«ҚР ҰҒА Хабарлары. Физика-математика сериясы».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы).

Ақпарат агенттігінің мерзімді баспасөз басылымын, ақпарат агенттігін және желілік басылымды қайта есепке қою туралы ҚР Мәдениет және Ақпарат министрлігі «Ақпарат комитеті» Республикалық мемлекеттік мекемесі **28.02.2025** ж. берген **№KZ20VPY00113741** Куәлік.

Тақырыптық бағыты: *ақпараттық-коммуникациялық технологиялар*

Қазіргі уақытта: *«ақпараттық-коммуникациялық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

«Известия НАН РК. Серия физико-математическая».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*

Свидетельство о постановке на переучет периодического печатного издания, информационного агентства и сетевого издания **№ KZ20VPY00113741.** Дата выдачи **28.02.2025**

Тематическая направленность: *информационно-коммуникационные технологии.*

В настоящее время: *вошел в список журналов, рекомендованных КОКСНВО МНВО РК по направлению «информационно-коммуникационные технологии».*

Периодичность: *4 раза в год.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

News of the National Academy of Sciences of the Republic of Kazakhstan.
Series of Physics and Mathematics
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty).
Certificate **No. KZ20VPY00113741** on the re-registration of the periodical printed and online publication of the information agency, issued on **28.02.2025** by the Republican State Institution «Information Committee» of the Ministry of Culture and Information of the Republic of Kazakhstan
Subject area*: information and communication technologies.*
Currently: *included in the list of journals recommended by the CCSES MSHE RK in the direction of «Information and communication technologies».*
Periodicity: *4 times a year.*
Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

NEWS OF THE NATIONAL ACADEMY OF SCIENCESOF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 1. Namber 353 (2025). 124–135

https://doi.org/10.32014/2025.2518-1726.329

IRSTI 28.23.37

A.N. Zhidebayeva [1], S.T. Akhmetova [2], A.O. Aliyeva[3], B.O. Tastanbekova[2], G.S. Shaimerdenova[2*], 2025.

[1]Peoples' Friendship University named after academician A. Kuatbekov, Shymkent, Kazakhstan;
[2]Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan;
[3]Khoja Akhmet Yassawi International Kazakh – Turkish University, Turkistan, Kazakhstan.
E-mail: danel101kz@gmail.com

REVIEW OF DETECTION AND PREVENTION OF OFFENSIVE LANGUAGE VIA SOCIAL MEDIA DATA MINING

Zhidebayeva Aziza – Candidate of Technical Sciences, Senior Lecturer, department of Computer science and Mathematics, Peoples Friendship University named after Academician A. Kuatbekov, Shymkent, Kazakhstan, E – mail: aziza_68.kz@mail.ru, https://orcid.org/0000-0002-3768-4835;
Akhmetova Sabira – candidate of physical and mathematical sciences, associate professor, department of Information system, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan, E – mail: sabdas65@mail.ru, https://orcid.org/0000-0001-5164-2028;
Aliyeva Akbayan – candidate of technical sciences, department of Computer engineering, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan, E – mail: aliyevaakbayan@mail.ru, https://orcid.org/0009-0006-6932-7302;
Tatanbekova Bayan – master, teacher, department of Information Communication technologies, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan, E – mail: baya2013@inbox.ru , https://orcid.org/ 0000-0002-4549-0344;
Shaimerdenova Guldana – PhD, associate professor, department of Information Communication technologies, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan, E – mail: danel101kz@gmail.com , https://orcid.org/0000-0001-8685-7125.

Abstract. This study delves into the multifaceted problem of cyberbullying among children and adolescents, focusing on developing an adapted framework to address this pervasive problem in the context of Kazakh social media. Based on a comprehensive analysis of international research, including contributions from the United States, England, and European countries, the paper integrates theoretical methodologies such as analysis, synthesis, empirical comparison, and experimentation. By contextualizing the findings in the unique cultural and digital landscape of Kazakhstan, the paper seeks to bridge the gap between global insights and local applications.

A central aspect of this study is the exploration of technical methodologies to combat cyberbullying. In particular, it examines the creation of a parser and systematic collection of data to train machine and deep learning algorithms capable of identifying and mitigating offensive or derogatory language in real time. These technologies promise to improve proactive moderation of online platforms and protect vulnerable users. Aimed at both novice and seasoned data scientists, this study highlights the need for further study into the practical implementation of deep learning algorithms. As part of the larger ongoing project, future work will refine these approaches, with subsequent publications looking in more detail at the deployment and optimization of advanced algorithmic models.

Keywords: classification, machine learning, deep learning, social media, social networks

А.Н. Жидебаева [1], С.Т. Ахметова [2], А.О. Алиева[3], Б.О. Тастанбекова[2], Г.С. Шаймерденова[2*], 2025.

[1]Академик Ә. Қуатбеков атындағы Халықтар достығы университеті, Шымкент, Қазақстан;

[2]М. Ауезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан;

[3]Қожа Ахмет Ясауи атындағы Халықаралық Қазақ-Түрік Университеті, Түркістан, Қазақстан.

E-mail: danel101kz@gmail.com

ӘЛЕУМЕТТІК ЖЕЛІЛЕРДЕН DATA MINING АРҚЫЛЫ БЕЙӘДЕП СӨЗДЕРДІ АНЫҚТАУ ЖӘНЕ АЛДЫН АЛУҒА ШОЛУ

Zhidebayeva Aziza – т.ғ.к., Академик А. Қуатбеков атындағы Халықтар достығы университетінің «Информатика және математика» кафедрасының аға оқытушысы, Шымкент, Қазақстан, E – mail: aziza_68.kz@mail.ru , https://orcid.org/0000-0002-3768-4835;

Akhmetova Sabira – физика-математика ғылымдарының кандидаты, доцент, Ақпараттық жүйе кафедрасы, Мұхтар Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан, E – mail: sabdas65@mail.ru, https://orcid.org/0000-0001-5164-2028;

Aliyeva Akbayan – техника ғылымдарының кандидаты, «Компьютерлік техника» кафедрасы, Қожа Ахмет Яссауи атындағы Халықаралық Қазақ-Түрік университеті, Түркістан, Қазақстан, E – mail: aliyevaakbayan@mail.ru, https://orcid.org/0009-0006-6932-7302;

Tatanbekova Bayan – магистр, оқытушы, ақпараттық коммуникациялық технологиялар кафедрасы, Мұхтар Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан, E – mail: baya2013@inbox.ru , https://orcid.org/ 0000-0002-4549-0344;

Shaimerdenova Guldana – Мұхтар Әуезов атындағы Оңтүстік Қазақстан университеті, «Ақпараттық коммуникациялық технологиялар» кафедрасының PhD, доценті, Шымкент, Қазақстан, E – mail: danel101kz@gmail.com, https://orcid.org/0000-0001-8685-7125.

Аннотация. Бұл зерттеу қазақстандық әлеуметтік медиа контекстінде кең таралған проблеманы шешу үшін бейімделген құрылымды әзірлеуге назар аудара отырып, балалар мен жасөспірімдер арасындағы кибербуллингтің көп қырлы проблемасын зерттейді. Америка Құрама Штаттарының, Англияның

және Еуропа елдерінің үлестерін қоса алғанда, халықаралық зерттеулерді жан-жақты талдау негізінде жұмыс талдау, синтез, эмпирикалық салыстыру және эксперимент сияқты теориялық әдіснамаларды біріктіреді. Қазақстанның бірегей мәдени және цифрлық ландшафтындағы қорытындыларды контекстке келтіру арқылы мақала жаһандық түсініктер мен жергілікті қолданбалар арасындағы алшақтықты жоюға тырысады.

Бұл зерттеудің орталық аспектісі кибербуллингпен күресудің техникалық әдістемелерін зерттеу болып табылады. Атап айтқанда, ол нақты уақыт режимінде қорлайтын немесе кемсітетін тілді анықтауға және жеңілдетуге қабілетті машиналық және терең оқыту алгоритмдерін үйрету үшін талдаушыны құруды және деректерді жүйелі жинауды зерттейді. Бұл технологиялар онлайн платформалардың белсенді модерациясын жақсартуға және осал пайдаланушыларды қорғауға ықпал етеді.

Жаңадан келген ғалымдарға да, тәжірибелі мамандарға да бағытталған бұл зерттеу терең оқыту алгоритмдерін практикалық енгізу бойынша әрі қарай зерттеу қажеттілігін көрсетеді. Ағымдағы үлкен жобаның бөлігі ретінде болашақ жұмыс осы тәсілдерді нақтылайды, ал кейінгі жарияланымдар кеңейтілген алгоритмдік үлгілерді қолдану және оңтайландыру мәселелерін толығырақ қарастырады.

Түйін сөздер: классификация, машиналық оқыту, терең оқыту, әлеуметтік медиа, әлеуметтік желілер

А.Н. Жидебаева[1], С.Т. Ахметова[2], А.О. Алиева[3], Б.О. Тастанбекова[2], Г.С. Шаймерденова[2*], 2025.

[1]Университет Дружбы Народов имени академика А. Куатбекова, Шымкент, Казахстан;

[2] Южно-Казахстанский университет имени Мухтара Ауэзова, Шымкент, Казахстан;

[3] Международный Казахско-Турецкий Университет им. Ходжа Ахмета Ясави, Туркестан, Казахстан.

E-mail: danel101kz@gmail.com

ОБЗОР ОБНАРУЖЕНИЯ И ПРЕДОТВРАЩЕНИЯ ОСКОРБИТЕЛЬНОЙ ЛЕКСИКИ С ПОМОЩЬЮ DATA MINING В СОЦИАЛЬНЫХ СЕТЯХ

Жидебаева Азиза – кандидат технических наук, старший преподаватель кафедры Информатики и математики, Университет Дружбы Народов имени академика А. Куатбекова, Шымкент, Казахстан, E-mail: aziza_68.kz@mail.ru, https://orcid.org/0000-0002-3768-4835;

Ахметова Сабира – кандидат физико-математических наук, доцент кафедры «Информационные системы» Южно-Казахстанского университета имени Мухтара Ауэзова, Шымкент, Казахстан, E-mail: sabdas65@mail.ru, https://orcid.org/0000-0001-5164-2028;

Алиева Акбаян – кандидат технических наук, кафедра «Вычислительная техника», Международный казахско-турецкий университет имени Ходжи Ахмета Ясави, Туркестан,

Казахстан, E-mail: aliyevaakbayan@mail.ru, https://orcid.org/0009-0006-6932-7302;
Тастанбекова Баян – магистр, преподаватель кафедры информационно-коммуникационных технологий, Южно-Казахстанский университет имени Мухтара Ауэзова, Шымкент, Казахстан, E-mail: baya2013@inbox.ru, https://orcid.org/ 0000-0002-4549-0344;
Шаймерденова Гульдана – PhD, доцент кафедры информационно-коммуникационных технологий Южно-Казахстанского университета имени Мухтара Ауэзова, Шымкент, Казахстан, E-mail: danel101kz@gmail.com, https://orcid.org/0000-0001-8685-7125.

Аннотация. Это исследование рассматривает многогранную проблему кибербуллинга среди детей и подростков, сосредотачиваясь на разработке адаптированной структуры для решения этой повсеместной проблемы в контексте казахских социальных сетей. Основываясь на всестороннем анализе международных исследований, включая вклад Соединенных Штатов, Англии и европейских стран, работа интегрирует теоретические методологии, такие как анализ, синтез, эмпирическое сравнение и экспериментирование. Контекстуализируя результаты в уникальном культурном и цифровом ландшафте Казахстана, статья стремится преодолеть разрыв между глобальными инсайтами и местными приложениями.

Центральным аспектом данного исследования является изучение технических методологий борьбы с кибербуллингом. В частности, рассматривается создание парсера и систематический сбор данных для обучения алгоритмов машинного и глубокого обучения, способных в реальном времени выявлять и смягчать оскорбительный или уничижительный язык. Эти технологии обещают улучшить проактивную модерацию онлайн-платформ и защитить уязвимых пользователей.

Нацеливаясь как на начинающих специалистов по обработке данных, так и на опытных профессионалов, это исследование подчеркивает необходимость дальнейшего изучения практической реализации алгоритмов глубокого обучения. В рамках более крупного текущего проекта будущая работа будет уточнять эти подходы, а последующие публикации будут более подробно рассматривать развертывание и оптимизацию продвинутых алгоритмических моделей.

Ключевые слова: классификация, машинное обучение, глубокое обучение, социальные медиа, социальные сети.

Introduction. Cyberbullying has become a growing social problem, especially during the pandemic, as traditional forms of mental and physical abuse have expanded to virtual environments. Although this form of aggression may initially seem harmless, it has significant differences compared to traditional bullying. The unique capabilities of the Internet – such as anonymity, the potential for a wide audience, 24/7 accessibility, and the ability to fake identities – have increased the impact and complexity of cyberbullying (Van,et al, 2018).

In 2017, a study conducted by the Russian Association of Electronic Communications (RAEC) identified cyberbullying as a pressing issue. This study

included structured surveys and independent analyses, although its results were partially disclosed to the public. The study involved 2,500 people aged 12 to 17 and 19 to 23, using targeted questions to examine the prevalence and characteristics of cyberbullying in these age groups. With the growing mental health risks of young people, the issue of online abuse is becoming increasingly important. Mental health, according to one concept, reflects "an individual's harmonious alignment with oneself and the environment, including other people, nature, and space." (Valle-Cano, et al, 2023).

Despite many surveys and analyses conducted on the topic, the development of methods to effectively detect or identify cyberbullying remains under-researched. For example, recent studies have begun to analyze specific text data from media platforms to identify patterns of online abuse (Van, et al, 2018). This shift is an important step toward automating the detection of cyberbullying, but significant challenges remain.

The prevalence of online bullying has increased alarmingly over the past two decades. Among US Internet users aged 10 to 17, instances of online bullying increased from 6% in 2000 to 9% in 2005, and then to 11% in 2010 (Khan, et al, 2022). These findings, from studies conducted by Finkelhor, Jones, and Mitchell (2013), highlight a clear trend of increasing online bullying. Similar trends have been noted in studies focusing on older adults, where 10% of respondents reported experiencing online bullying at least monthly. Interestingly, when researchers looked more deeply into hateful or harmful online content, these rates tended to be even higher, suggesting that cyberbullying is often underestimated in superficial surveys.

The findings are supported by a US review article, "The Changing Landscape of Peer Aggression: A Review of the Cyberbullying and Intervention Literature," which provides a comprehensive overview of the evolution of cyberbullying and the decline of traditional forms of bullying. While this study offers valuable insights, it relied solely on manual methods of data collection and analysis. Specifically, the researchers collected data through survey responses, manually organized and analyzed the information, and then grouped the results accordingly. This approach, while effective for descriptive research, highlights the lack of automated systems or software solutions to facilitate the detection and analysis of cyberbullying content. To address these gaps, recent research has begun to focus on technological interventions, such as the use of machine learning and natural language processing to detect offensive content in real time. These tools can potentially provide scalable solutions for monitoring and mitigating cyberbullying on digital platforms. However, the application of such tools is still in its infancy, requiring significant refinement and testing to account for the nuances of language, context, and cultural sensitivity (Kumari, et al, 2020). The cyberbullying is a complex and rapidly evolving problem, exacerbated by the unique characteristics of digital environments. While traditional survey research has provided basic insights, there is a pressing need for automated and scalable solutions to effectively combat online aggression. Future research

should prioritize the integration of advanced computational methods to identify, analyze, and mitigate cyberbullying, with an emphasis on protecting vulnerable populations and promoting mental well-being in the digital age (Elzayady, et al, 2023).

Materials and methodology

In general, all of these works shared the same central idea, but different technologies were employed according to the goal that was more precisely defined. In addition to using machine and deep learning algorithms, some areas of the world involved more work with people, such as polling and pattern recognition. Databases play a critical role in studies where machine learning algorithms are mainly applied, and this is where the research methodology is usually divided into two main stages. In the first stage, datasets are often taken from public resources such as social media platforms, open access databases, or pre-existing repositories. The second stage involves the creation of customized data collection tools specifically designed to achieve the research goals. These tools are often accompanied by manual data labeling and cleaning to prepare them for further analysis. Once cleaned, the data is processed and experiments are conducted with various machine learning algorithms to detect and classify patterns such as cyberbullying. Unlike traditional approaches, our study stands out because it not only models harassment "attacks" but also examines the broader context of cyberbullying, including the more subtle interactions between perpetrators, victims, and witnesses. This broader approach is especially important because interactions between these three groups often signal that cyberbullying is occurring or escalating. Thus, rather than focusing on individual instances of harmful behavior, this study aims to identify the full spectrum of cyberbullying indicators, which may be critical to identifying ongoing harassment (Ghozali, et al, 2023). Despite the progress made, a fully automated system for real-time cyberbullying detection does not yet exist, and current systems still require human intervention to check when potential indicators of cyberbullying emerge. One of the challenges faced by previous studies is the uneven distribution of data, where only 4-7% of detected cases were labeled as positive for cyberbullying, and the majority of data was neutral or irrelevant (Jin-Liang, et al, 2018).

In contrast, (Agrawal, et al, 2018) conducted a study using Twitter, where they applied a binary classification approach to detect cyberbullying. They created a system capable of extracting the semantic meaning of words in tweets, using these word-level features along with additional text features such as sentiment analysis and vocabulary usage to differentiate between positive and negative sentiment. Their approach included both a binary classification model to determine the presence or absence of cyberbullying and a multi-class classification model to categorize the severity of the behavior as low, medium, or high. The methodology used in this study highlights the potential for more nuanced and scalable approaches to detecting cyberbullying on online platforms. However, challenges such as imbalances in datasets and the need for manual validation remain, requiring further improvements in algorithmic accuracy and data quality (Rezvan, et al, 2018).

The Random Forest classifier, combined with SMOTE, demonstrated excellent results with a kappa coefficient of 0.711, an overall accuracy of 91.153%, and an f-score of 0.898, achieving the highest overall classifier performance. In the binary classification setting, it performed even better with an AUC of 0.971 and an f-score of 0.929, confirming its effectiveness. The significance of the proposed features was highlighted through a comparative analysis with standard features in both multi-class and binary classification problems. The study highlights the need to explore a dataset that includes a balanced distribution of neutral and cyberbullying posts to enhance understanding and further improve the classification framework (Talpur, et al, 2020).

Table 1. Analysis of compared articles

Reference	Algorithms Used	Metrics	Corpus Size (words)	Application Area	Year	Notable Observations
[1]	Binary classification, Linear Regression	F-score, Precision, Recall, AUC, MSE	85,462	Sentiment Analysis	2019	Demonstrated high AUC for detecting sentiment polarity.
[3]	Naïve Bayes, KNN, Decision Trees, Random Forest, SVM	F-score, Precision, Recall, AUC	24,189	Cyberbullying Detection	2018	Random Forest outperformed others with the highest F-score and precision.
[14]	Logistic Regression, SGD Classifier, Random Forest, SVM, LGBM Classifier	Accuracy, Recall, Precision, F1 Score	37,373	Fake News Detection	2021	LGBM Classifier achieved the highest F1 score for binary classification tasks.
[12]	LDA, SVM, Linear Regression, Multiple Classification, Deep Learning	Accuracy, Recall, Precision	15,874	Topic Modeling	2020	Deep Learning models were more robust for complex datasets compared to traditional classifiers.
[6]	Clustering, Classification, Dimensionality Reduction	Accuracy, Recall, Precision	5,000	Spam Detection	2017	Clustering provided meaningful insights into unlabeled data, aiding classification tasks.
[17]	LDA, Random Forest, KNN, Classification, and Multiple Classification	Accuracy, Recall, Precision, MSA, MAE	17,546	Text Categorization	2022	Combined classifiers improved performance; MSA and MAE highlighted model efficiency for specific categories.

The methodologies employed in these studies demonstrate a consistent approach that emphasizes systematic data collection and preprocessing. Figure 1 illustrates the process of extracting information from a rich media landscape, highlighting

two primary scraping methods: API-based and custom system designs. API-based scraping requires token authentication, allowing for seamless and efficient access to structured data. In contrast, custom scraping leverages user agents to extract data from web resources, allowing for greater flexibility in analyzing system architecture. The choice between these methods depends on the requirements of the task—API integration offers speed and convenience, while custom scraping supports more complex, customized data collection needs (Masadeh, et al, 2022). This dual approach provides adaptability by enabling complex data preprocessing, which is critical for robustly training machine learning models and optimizing system design.

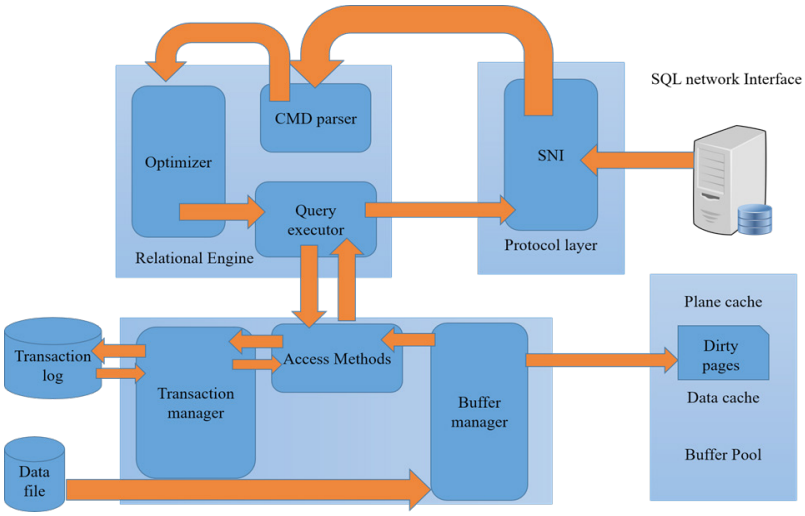

Fig. 1. Data processing scheme for web resources

Kaggle.com is an indispensable platform for newcomers to the data science community, offering an accessible entry point to learn and develop skills in the field. Users can upload their datasets and participate in competitions that require building predictive models, providing a great opportunity to work with real data. Although a wide range of datasets are available for free, the lack of an appropriate dataset for our specific research topic remains a challenge for this study (Sultan, et al, 2021).

Machine learning uses a variety of algorithms for classification and prediction problems. Logistic regression, a fundamental statistical model, is widely used for binary classification problems. It works by modeling the probability of a binary outcome as a function of one or more predictor variables, often using a logistic (sigmoid) function to produce output values between 0 and 1. Logit regression, a variant of logistic regression, focuses on analyzing the limits or boundaries of a classification model, providing valuable insights into how the predictor variables influence the outcome.

Introduced in 1951 by Evelyn Fix and Joseph Hodges, K-Nearest Neighbors (KNN) is a nonparametric machine learning method used primarily for classification problems, although it can also be applied to regression. KNN classifies a data point based on the majority of its nearest neighbors in the feature space. The method makes no assumptions about the distribution of the data, making it very flexible and effective for a wide range of problems. However, its performance can degrade when dealing with high-dimensional data due to the "curse of dimensionality"(Comito, et al, 2019).

Linear regression, another basic machine learning technique, models the relationship between a scalar dependent variable and one or more independent variables. Simple linear regression involves a single independent variable, while multiple linear regression can handle multiple predictors. This technique is widely used in economics, health care, and social sciences to understand the influence of various factors on an outcome. Its versatility in predicting continuous outcomes makes it a key tool in predictive analytics.

These algorithms, although varying in complexity, are critical to building accurate models in a variety of fields, such as finance, health care, and social media analytics.The SVM (Supper Vector Machine) is a controlled learning model that analyzes data for order and relapse investigation. In addition to conducting direct grouping, SVMs can effectively use a technique known as the part stunt to play out a nonstraight characterisation, verifiably planning their contributions to high-dimensional element spaces.

Random Forests are a group learning technique for characterization, relapse, and various assignments that produce the class that is the method of the classes (order) or mean/normal forecast (relapse) of the individual trees by developing a large number of decision trees at preparation time. Choice trees have a tendency to overfit their preparation set in irregular choice woodlands. Although irregular woodlands typically encircle preferred trees, their accuracy is inferior to that of angled-helped trees. In every situation, the presentation of the information can be affected by its properties.

Make comparisons between the algorithms.

The analysis and presentation of the authors' findings on this topic was all fairly strong. Most often used metrics to assess the algorithms were:

$$Pr = \frac{TP}{TP+FP} \quad (1)$$

Precision (Pr): Precision is the ratio of true positives (TP) to the total number of predicted positives, which includes both true positives (TP) and false positives (FP). This reflects how many of the objects that the model classified as positive are actually positive. Precision is important in scenarios where false positives are costly or undesirable.

$$recall = \frac{TP}{TP+FN} \quad (2)$$

Recall (or Sensitivity): Recall is the ratio of true positives (TP) to the total number of true positives, which includes both true positives (TP) and false negatives (FN). It measures the ability of the model to correctly identify all relevant cases in the dataset. Recall is especially important in contexts where false negatives are more critical, such as in medical diagnostics.

$$F1 = \frac{2\,precision*recall}{precision+recall} \quad (3)$$

F1 Score: The F1 score is the harmonic mean of precision and recall, providing a balanced measure between the two. It is particularly useful when there is a need to balance precision and recall, and is often used when the class distribution is unbalanced.

$$accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (4)$$

Accuracy: Accuracy is the ratio of correctly predicted cases (both true positives and true negatives) to the total number of cases in the dataset. While accuracy is easy to understand, it can be misleading in imbalanced datasets where the number of negative examples greatly outnumbers the positive ones (Landau, et al, 2022).

These metrics are critical to understanding how well a classification model performs in various aspects, such as identifying true positives, avoiding false positives, and minimizing false negatives. Depending on the context, you may want to prioritize precision, recall, or F1-score over accuracy.

When determining whether the model operates in a relevant manner, those four marks are the most frequently used.

Results and discussion

There are several related research studies on this subject, one of which is mentioned in the introduction. This piece focuses on a few key areas of recent writing on cyberbullying among young people nowadays. We specifically need a better understanding of which young people are most likely to be threats, victims, or observers online, as well as the circumstances that will lead young people to speak out against cyberbullying behavior. The characteristics of current anti-cyberbullying initiatives being used in American schools and the gathering of young people for these activities should be noted as an important component of our investigation. This information will make it easier to decide which mediations need to be strengthened and expanded and which need to be replaced.

Utilizing artificial intelligence to find hate speech in text. The results of applying multiple machine learning techniques to classify hate speech using different feature combinations are compared in this section. The following strategies are among the

most widely used in modern research for building and training classifiers: Picture of decision tree, random forest, naive Bayes, logistic regression, k-nearest neighbors, and support vector machine

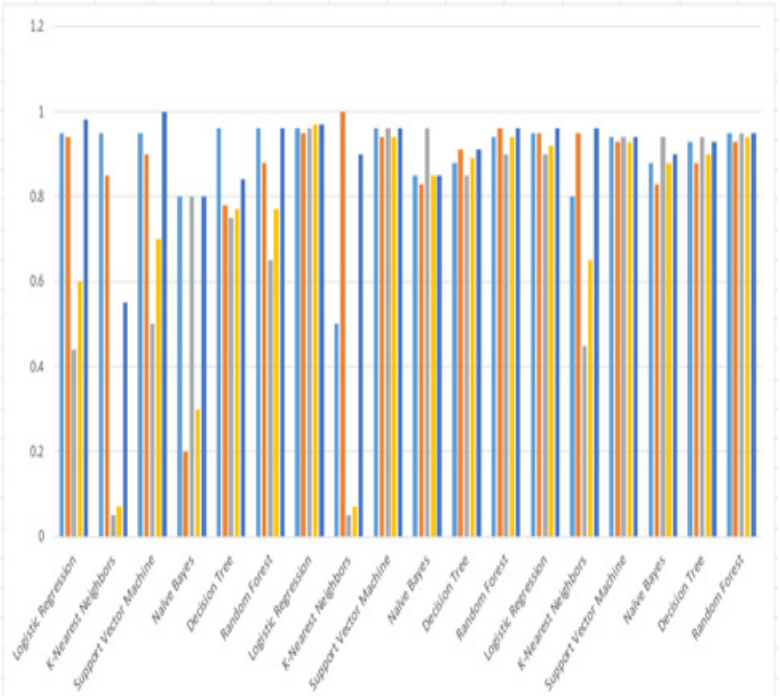

Fig. 2. Comparison of results of classification algorithms

The study synthesizes key findings related to the conceptualization of cyberbullying, making comparisons with traditional bullying, highlighting their common features and varying degrees of severity. Different types of cyberbullying are explored, profiling the characteristics of both cyberbullies and their victims. The impact of cyberbullying on the psychosocial development of young people is also considered, taking into account the factors of age and sexual orientation. In addition, the study delves into the underlying causes of cyberbullying and describes intervention strategies at the individual, organizational and societal levels. Recommendations for young people, parents, educators and schools are presented to effectively address and mitigate the problem from a holistic well-being perspective.

Conclusion. Based on the findings of previous studies, we decided to automate the process of detecting cyberbullying in text data. Our first step was to investigate whether similar studies had been conducted in Kazakhstan or in the former Soviet Union. Although we found some conceptually related works, none of them fully met our specific requirements for the project. This led us to the conclusion that we needed to develop our own autonomous system capable of automatically detecting

cases of cyberbullying and implementing preventive measures. Given the lack of existing tools that could meet our needs, we decided to take on this task ourselves, focusing on creating a customized solution to fill this gap. This approach allowed us to tailor the system to the unique needs of our study and the specific characteristics of the Kazakh digital environment.

References

Van Hee C., Jacobs G., Emmery C., Desmet B., Lefever E., Verhoeven B., ... & Hoste V. (2018). Automatic detection of cyberbullying in social media text. PloS one, 13(10), e0203794. (in English)

G. Valle-Cano L. Quijano-Sánchez F. Liberatore, and J. Gómez. "Social haterBERT: a dichotomous approach for automatically detecting hate speech on twitter through textual analysis and user profiles," Expert Systems with Applications, vol. 216, no. 1, pp. 1-17, 2023. (in English)

S. Khan, M. Fazil, V. Sejwal, M. Alshara, R. Alotaibi et al., "BiCHAT: BiLSTM with deep CNN and hierarchical attention for hate speech detection," Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 7, pp. 4335-4344, 2022. (in English)

Kumari K., Singh J. P., Dwivedi Y.K., & Rana N.P. (2020). Towards Cyberbullying-free social media in smart cities: a unified multi-modal approach. Soft Computing, 24(15), 11059-11070. (in English)

H. Elzayady, M. Mohamed, K. Badran and G. Salama, "A hybrid approach based on personality traits for hate speech detection in Arabic social media," International Journal of Electrical and Computer Engineering, vol. 13, no. 2, pp. 1979-1988, 2023. (in English)

I. Ghozali, K. Sungkono, R .Sarno and R. Abdullah, "Synonym based feature expansion for Indonesian hate speech detection," International Journal of Electrical & Computer Engineering vol. 13, no. 1, pp. 2088-8708, 2023. (in English)

Jin-Liang Wanga, Linda A. Jackson , James Gaskin , Hai-Zhen Wang, "The effects of Social Networking Site (SNS) use on college student's friendship and wellbeing", Elsevier, Computers in Human Behavior 37 (2014). (in English)

Agrawal S., & Awekar A. (2018, March). Deep learning for detecting cyberbullying across multiple social media platforms. In European conference on information retrieval (pp. 141-153). Cham: Springer International Publishing. (in English)

Rezvan M., Shekarpour S., Balasuriya L., Thirunarayan, K., Shalin, V. L., & Sheth, A. (2018, May). A quality type-aware annotated corpus and lexicon for harassment research. In Proceedings of the 10th acm conference on web science (pp. 33-36). (in English)

Talpur B.A., & O'Sullivan D. (2020). Cyberbullying severity detection: A machine learning approach. PloS one, 15(10), e0240924. (in English)

M. Masadeh, H. Davanager, and A. Muaad, "A novel machine learning-based framework for detecting religious Arabic hatred speech in social networks," International Journal of Advanced Computer Science and Applications, vol. 13, no. 9, pp. 767-776, 2022. (in English)

Sultan D., Suliman A., Toktarova A., Omarov B., Mamikov S., & Beissenova G. (2021, January). Cyberbullying detection and prevention: Data mining in social media. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 338-342). IEEE. (in English)

C. Comito, A. Forestiero and C. Pizzuti, "Word embedding based clustering to detect topics in social media," In 2019 IEEE/WIC/ACM Int. Conf. Web Intell. WI 2019, Thessaloniki, Greece, pp. 192–199, 2019 (in English)

A. Landau, S. Ferrarello, A. Blanchard, K. Cato, N. Atkins et al., "Developing machine learning-based models to help identify child abuse and neglect: key ethical challenges and recommended solutions," Journal of the American Medical Informatics Association, vol. 29 no. 3, pp. 576-580, 2022. (in English)

CONTENTS

INFORMATION AND COMMUNICATION TECHNOLOGIES

МАЗМҰНЫ

АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР

СОДЕРЖАНИЕ

ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

Publication Ethics and Publication Malpracticein the journals of the National Academy of Sciences of the Republic of Kazakhstan

For information on Ethics in publishing and Ethicalguidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any otherlanguage, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:
www:nauka-nanrk.kz
http://physics-mathematics.kz/index.php/en/archive
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Директор отдела издания научных журналов НАН РК *А. Ботанқызы*
Редакторы: *Д.С. Аленов, Ж.Ш.Әден*
Верстка на компьютере *Г.Д.Жадыранова*

Подписано в печать 20.03.2025.
Формат 60х881/8. Бумага офсетная. Печать –ризограф.
20,0 п.л. Заказ 1.

Национальная академия наук РК
050010, Алматы, ул. Шевченко, 28, т. 272-13-19