DEPRESSION DETECTION IN SOCIAL MEDIA USING BAGGING CLASSIFIER

Article *in* Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition) · January 2023

DOI: 10.17605/OSF.IO/EVXB7

Muhammad Javaid Iqbal
Superior University

Arfan Jaffar
Superior University

Sobia Yaqoob
University of Okara

Inam Ul Haq
University of Okara

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

DEPRESSION DETECTION IN SOCIAL MEDIA USING BAGGING CLASSIFIER

FATIMA NOOR
Department of Computer Science, Superior University Lahore, Pakistan.
Email: mscs-f20-004@superior.edu.pk

MUHAMMAD JAVAID IQBAL*
Department of Computer Science, Superior University Lahore, Pakistan.
Corresponding Author Email: javaid.iqbal@supeior.edu.pk

SOBIA YAQOOB
Lecturer, Department of Computer Science, University of Okara, Pakistan.
Email: sobia.yaqoob@uo.edu.pk

SHAHID MEHMOOD
Department of Computer Science, University of Central Punjab, Lahore, Pakistan.
Email: shahid.mehmood01@ucp.edu.pk

ARFAN JAFFAR
Ph.D. Professor, Dean Department of Computer Science and Information Technology, the Superior University Lahore, Pakistan. Email: arfan.jaffar@superior.edu.pk

INAM UL HAQ
PhD Scholar, Department of Information Technology, University of Okara 56300 Pakistan.
Email: inamulhaq@uo.edu.pk

Abstract

We live in such a world where almost everybody has a lot of problems like discomfort and tension in their lives irrespective of age, gender, or cast. People are struck with inferiority, others take the pressure, and a few become hopeless and try to destroy themselves. The rapid development of technology made them feel comfortable sharing their feelings on social media. Web-based life has, as of late, developed as a chief technique to scatter data on the web. It has been created as a unique point for a considerable number of people to impart their musings, individual encounters, and social goals with intrigued companions and offer assessments, photographs, and recordings mirroring their emotions. This gives a chance to examine informal community information for people’s emotions to explore their dispositions and frames of mind when conveying through these online instruments. We applied the Bagging classifier over various informational collections utilizing a Bag of words chosen beforehand to perform probes of different informational indexes. We found that Bagging Classifier gives better and more promising outcomes when contrasted with the existing classifier. Packing exploits gathering learning wherein numerous powerless students beat a solitary solid student. It diminishes fluctuation and encourages us to stay away from overfitting. By utilizing a corpus of 3.5M tweets, we accomplished a 91% exactness rate in the grouping, with an accuracy score of 86. We accept that this strategy might help create instruments that gauge the danger of an individual being discouraged and can be utilized by doctors, concerned people, and medicinal services organizations to help determine.

Keywords: Deression Detection, Machine Learning, Bagging Classifier, Twitter Analysis

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

1. INTRODUCTION

Teams that work on software development and operations usually divide their tasks. These silos may have reduced the risks to production settings, but they are also known to slow down code modifications and deployment significantly. As operations work to ensure software's stability and performance in production, tensions frequently occur as developers try to satisfy customer requests by releasing new software releases. We have explained the Software Project's failure without using DevOps and Software Project's success with DevOps in Table 3. The business, however, requires more regular releases and cannot afford the process of putting software into production to be time-consuming, demanding, and prone to mistakes [1]. DevOps's primary goal is to promote an agile software development lifecycle using continuous development techniques, including delivery, employment, and microservices [2].

The agile approach (2001) raised the bar for conventional methodologies, and the introduction of the DevOps concept in 2009 marked a further significant advancement (Development Operations) [3]. DevOps is a growing concept that will likely become essential for staying competitive. The DevOps idea is being adopted by more and more IT companies [4]. However, many firms are unsure of the specific talents required to complete the many responsibilities expected of a DevOps team. One of the main reasons IT projects fail is a lack of expertise in the project environment [5]. DevOps in an industry necessitates the integration of various essential components that were [6] once thought of as entirely separate processes in the software distribution process. I refer to the fundamental elements of DevOps as critical to its efficient operation, without which it cannot function [6] [7].

The DevOps Trend is concentrated on facilitating client interaction and cooperation, which is essential to providing software. The client frequently receives new software features from businesses that offer software and worldwide web solutions [8]. For companies, this constant software deployment has presented equal benefits and difficulties. DevOps is a recent development that aids in this [9]. DevOps addresses the issue by bridging the gap between development and operations staff, who are in charge of deploying, managing, and supporting the systems at the client site [10] [11]. However, the description of the concurrent implementation of DevOps and Agile in companies has a research problem. The insufficient amount of research currently accessible in this area, the majority of which gives different perspectives on their implementation, prevents a properly thorough definition of the advantages of their coupled deployment [12]. Figure 1: DevOps: 3rd generation software development method.

As of late, Internet-based life has developed as a head strategy to spread data on the web and cloud. Through these online systems, many people impart their contemplation, individual encounters and social standards[1]. We utilize a publicly supported technique to incorporate a rundown of Twitter clients who are determined to have misery. Utilizing something like a time of earlier web-based social networking postings, we use a Bag of Words to evaluate each tweet [2]. Untreated despondency builds the opportunity for dangerous practices. The massive test of recognizing discouragement is acknowledging

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

that burdensome side effects may contrast with patients' conduct and character. For centre wretchedness, specialists may assess the patient using the downturn test taken by patients. These clinical records are confined because of numerous variables, for example, age and sex; also, they are private and costly. To conquer such confinements of clinical information, it is advantageous to utilize content mining devices to separate and examine melancholy indications from internet-based life, for example, Twitter. Web-based life creates endless information consistently in light of a large number of dynamic clients share and convey in the whole network. It changes human communication. Other than the customary information, for example, written works, internet-based life information is more extravagant and increasingly available. Exploring this new quick development of information requires propelled improvement apparatus to find valuable knowledge. These trend-setting innovations incorporate Natural Language Processing (NLP), information mining, AI, web-based life examination, etc. In our examination, the objective is to separate and condense the extraordinary yet possibly supportive components that burdensome indication performed from the web-based life information. At long last, the separated sorrow indications will be utilized as references when physically perceiving the clinical despondency by people [3].

Pakistan In recent years, [4]The pilot undertaking challenge contains performing early risk revelation of hopelessness by analyzing customer-made substance from Reedit. Towards this goal, a system gets customer-created content as data and should yield a desire regarding the customer's powerlessness to wretchedness. The pilot task informational index contains a customer-made substance, which is made and arranged consecutively. This considers watching the customer's progress and perceiving danger as on schedule as possible.

Customers are requested as a peril or non-threat Wretchedness). Each customer conveyed a progression of Reddit posts created within a given period. The pilot task was dealt with in two stages: getting ready and testing, each having a substitute dataset divided into ten pieces. During the getting ready location, a dataset containing a progressive plan of posts per customer was outfitted nearby the customer's class. All arrangement pieces were made available, including the all-out customer post-gathering.

During the test mastermind, the dataset of test customers was released sequentially (one release each week). Each release contained a bit of the customer post gathering, identifying with one bump (from the most settled to the freshest posts). Part systems are expected to yield gauges for customers subject to all present test pieces before the appearance of another bump. The meters could be either the characterization of a customer or no decision, up to the latest multi-day stretch of the test sort out where all customers must be given a class. Starting now and into the foreseeable future, we depict our gauge structure as reliant on an outfit-gathering approach, which merges oversaw learning information recovery, and feature assurance techniques.

As indicated by the Canadian Mental Health Association (2016), 20% of Canadians having a place with various socioeconomics have encountered dysfunctional behaviours during their lifetime, and around 8% of grown-ups have experienced significant melancholy. As indicated by World Health Organization (2014) measurements, about

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

20% of youngsters have experienced psychological maladjustments, and half of these behavior problems begin before age. Also, Approximately 23% of deaths worldwide were due to mental health and drug-related problems. The extent of suicide in Canada, where over 4,000 people have died by suicide and 90% of them were identified as having a psychiatric disorder, demonstrates the broad implications of dysfunctional conduct (Mental Health Commission of Canada, 2016). In addition to the severity of the psychiatric condition and its effects on a person's health and wellbeing, social stigma or segregation has led to people being ignored by the network or refusing to take necessary medications. This article investigates a few selected, profound neural system structures to recognise mental illness, specifically melancholy. We used the data released for the 2015 Computational Linguistics and Clinical Psychology (CLPsych) collaborative project (Coppersmith et al., 2015b). Our primary goal was to identify despondency using the best deep neural design from two of the most popular deep learning approaches in the field of regular language handling: Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), given the limited resources available. The assignment included three subtasks: recognising Post-Traumatic Stress Disorder (PTSD) versus control, despondency control, and PTSD versus discouragement [5].

Psychological wellness is and keeps on being a noticeable plague for the enlightened world. It is assessed that one out of four American residents experiences and analyzes mental issues at whatever year. When joined with the 2015 US Census for Residents 18 and more seasoned, these insights make an image of 80 million enduring United States natives. One of every three of these natives who experience the ill effects of dysfunctional behaviour may experience the ill effects of clinical sadness, propelling an abundance of concentration to handle this issue. From this substantive field, we centre on Major Depressive Disorder, regularly alluded to as clinical melancholy. Not exclusively do almost 300 million individuals overall experience the ill effects of clinical misery. However, the likelihood for a person to experience a significant burdensome scene in one year is 3 – 5% for guys and 8 – 10% for females. However, these sorrowful impacts are more remote than just cultural joy: Depression negatively affects U.S. organizations that add up to over $70 billion yearly lost in medicinal uses, efficiency, and comparative expenses. An extra $23 billion in different costs may accumulate in the interest of an individual, therefore influencing workdays, reducing work propensities, and conceivably inducing entanglements with focus, memory, and basic leadership practices. Our workplaces are another technique for building our classifier by regarding society as a book arrangement issue instead of a conduct one via web-based networking media stages. By utilizing a corpus of 2.5M tweets, we accomplished an 81% exactness rate in characterization, with an accuracy score of .86. We accept that this technique might help create devices that gauge the danger of an individual being discouraged, can be utilized by doctors, concerned people, and human services offices to assist in finding, even perhaps empowering those experiencing gloom to be progressively proactive about recuperating from their emotional wellness[2].

Informal organizations have already been established as fantastic platforms for customers to interact with their curious friends and share their observations, pictures, and recordings that reflect their moods, feelings, and evaluations. This allows one to look through unofficial community data on clients' emotions and ideas to explore their mindsets and demeanours as they interact through such internet tools. The proliferation of online informal groups and web and correspondence technologies, remarkably, has restored how people communicate and interact online. Applications like Facebook, Twitter, Instagram, and the like enable customers to share their emotions, thoughts, and opinions on a subject, issue, or topic online, in addition to hosting written and mixed-media content. On the one hand, this is fantastic for users of long-distance informal communication sites to directly and unapologetically contribute to and respond to any point on the web; from the other side, it offers incentives for those involved in the care to understand what may be occurring at the psychological illness of one who responded to a theme in a particular way[1].

The World Health Organization (WHO) predicts that 800,000 people worldwide attempt suicide by suicide annually, with an equal number also attempting suicide. The pain experienced as a result of such an incident is made worse by the possibility of suicide. This reality of suicide has convinced WHO member states to commit to reducing the rate of suicide by a significant percentage by 2020. The American Foundation for Suicide Prevention (AFSP) has identified characteristics or circumstances that may increase a person's risk to educate the general public. The following are the three main hazards: 1) Well-being elements (such as mental health and unrelenting suffering), 2) Ecological factors (such as provocation and unpleasant life events), and 3) Verifiable variables (e.g., past suicide endeavours and family history).

Additionally, the period before suicide may include clues about a person's struggle. Orders the following warning signs when pursuing: 1) speech (e.g., talking about being weight or lacking the will to live), 2) behaviour (e.g., refraining from exercising or sleeping excessively or insufficiently), and 3) mental state (e.g., melancholy or wrath). The first step towards avoiding suicide is realizing these risk factors.

In any case, the scandalous social situation and mental diseases imply that in danger, people may maintain a strategic distance from proficient help. They might be more ready to go to less conventional assets for use. Online web-based social networking platforms have recently shown to be one such sporadic asset. According to research, people in danger turn to modern innovations (such as meetings or small-scale online diaries) to communicate their most intense conflicts without directly facing anybody. In a new discipline, suicide risk factors and warning signals have already been discovered. On Twitter, Facebook, and other social media sites, there have even been instances of persons who committed suicide being used to write their final thoughts. We acknowledge that this vast amount of data on people's attitudes and behaviours may efficiently identify behavioural changes in persons in danger and may even be used to forecast death. This topic has received much attention from social figure research in recent years. However, a few activities have expressed concern over the lunchtime Twitter discovery of self-destructive thoughts. Current discovery tactics heavily rely on physically observed speech, which can limit their effectiveness partly because of the changing sorts of suicide notice indications in vulnerable people. Many of these

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

methods also focus on the messages that individuals at a specific moment spread independently of the environment, which may be addressed by the timing of products in the future[6].

The World Health Organization determined that the recession is fully accountable for even more than 4% of the Disability-Adjusted Life Years (DALYs) lost and, if the trend continues, would overtake ischemic heart disease as the second leading cause of DALYs lost by 2020. A total of 350 million people worldwide experience depression, which accounts for 11.9% of all years lived with disabilities (YLDs) and is the most prevalent of all psychiatric and neurological diseases. In 2000, wretchedness forced a yearly monetary weight of 83 billion dollars in the US - a large portion of which was credited to decrease efficiency and expand therapeutic costs. Discouragement is additionally a significant reason for suicide: as per an investigation by Goodwin and Jamison, 15-20% of all critical burdensome issues patients take their lives. This result is, to a great extent, avoidable if there are b appropriate intercessions; what's more, early recognition of wretchedness is the initial move towards this mediation. Most investigations of early recognition of wretchedness depend on determinations dependent on patients' self-detailed encounters and studies. These results are costly; as of 2009, only 30% of world countries that provide essential social insurance administrations have these initiatives. The omnipresence of internet-based life among the total populace can answer this issue. Studies have demonstrated affiliations between the use of online life and gloom. Exercises in web-based life can be utilized as indicators for prosperity and social interest. In any case, the immediacy of identification is crucial in seeing hopelessness on social media: the more we wait there to step in, the more prominent the danger of self-mischief. Thus, anticipating discouragement early in a client's life cycle is fundamental. This contends for models that don't simply look at one depiction of a client's exercises but rather track the client's exercises after some time. It also contends for assessment measurements that consider the exactness and review of identifying discouraged clients and the speed of that discovery[7].

This research study aims to study intelligent techniques and their usage, investigate depressed users using social media data sets, and identify the best classifier to improve results, evaluate the relative research effort of Depression detection in data science, and provide a solution for cross-culture users.

People have worked on depression detection in social media users of specific cultures, races and countries. All research is done using Text-based data, but we will analyse using text and emoji. Emojis are essential and express thoughts, vision and emotions better than textual data. Our primary purpose and objective are to identify depression-affected users early to diagnose and control it. There is no research found on the cross or multi-culture-oriented solution for depressed people. Our main objective is to provide a solution for cross-culture depressed users.

We analyze the scientific research productivity and intra and international collaborations of Social Media because we do not find any comprehensive study conducted to measure and identify depression among people using social media. Social media data under machine learning techniques to promote research and international collaboration

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

among people. Secondly, we introduce a new way to calculate depression in social media users using textual and emoji data. Our primary purpose is to measure this viral and rapidly developing disease early to control and overcome the suicidal issue. Many people suicide due to depression, inferiority and complexity. Every second, millions of users share data over the internet, including secrets, emotions, frustration and exploring ideas. Even all events and activities are being recorded over the cloud. Most of the time, we share things over social media that are not shared with our closest and physical relations. We have to use this data positively and diagnose depression early on before it grows to the final stages and takes someone's life. To calculate this, we use machine learning techniques and artificial intelligent methods.

2. LITERATURE REVIEW

Tiancheng Shen et al. [8] recommended that Depression location is a huge issue for human prosperity. In past investigations, online identification has demonstrated viability in Twitter, empowering proactive consideration for discouraged clients. Inferable from social contrasts, recreating the technique to other web-based life stages, for example, Chinese Weibo, in any case, may prompt horrible showing in light of deficient accessible named (self-announced sorrow) information for model preparation. Throughout this article, we study an interesting, testing issue of improving recognition in a specific objective space (for example, Weibo) with sufficient Twitter information as the reference space. We initially deliberately examine the downturn-related element designs crosswise over areas and abridge two significant recognition challenges: isomerism and divergence. The crucial data is moved crosswise over diverse regions using the Deep Neural Network model with Feature Adaptive Transformation and Combination system (DNN-FATC) that we also offer. Analyses exhibit improved execution contrasted with existing heterogeneous exchange strategies or preparing straightforwardly in the objective area (over 3.4% improvement in F1), showing the capability of our model to empower sadness identification through online life for more nations with various social settings. In this paper, we raised the issue of upgrading gloom identification through internet-based life with multi-source datasets. We proposed a cross-area Deep Neural Network model with Feature Adaptive Transformation and Combination procedure (DNN-FATC) to move the pertinent data crosswise over heterogeneous spaces. Exploratory outcomes checked the adequacy of our technique. Later on, we hope to further improve online location by consolidating disconnected investigates and adding to more individuals' prosperity.

Guangyao Shen et al. [9] expressed that Depression is a significant supporter of the general worldwide weight of ailments. Customarily, specialists analyze discouraged individuals up close and personal using alluding to clinical sadness criteria. In any case, over 70% of the patients would not counsel specialists at beginning periods of sorrow, which prompts further disintegration of their conditions. In the interim, individuals progressively depend on web-based networking media to unveil feelings and share their everyday lives. Subsequently, online networking has effectively distinguished physical and mental disorders. Motivated by such, our job anticipates making advantageous sadness discoveries through reaping online networking information. We build well-

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

named misery and non-sadness data - set on Twitter and concentrate on six gloom-related component gatherings encompassing the medical wretchedness requirements, yet in addition, online practices via web-based networking media. With these component gatherings, we propose a multi-modal, burdensome lexicon training algorithm to recognize the discouraged clients on Twitter. A progression of trials led to the approval of this model, which beats (+3% to +10%) a few benchmarks. We break down an extreme set of data on Twitter to uncover the hidden internet practices among discouraged and non-discouraged clients. This paper means to make auspicious sadness discovery through reaping web-based life. With the benchmark gloom and non-melancholy datasets just as well-characterized discriminative sorrow situated element gatherings, we proposed a multimodal burdensome lexicon learning strategy to identify discouraged clients on Twitter. We, at that point, examined the commitment of the component modalities and recognized discouraged clients on a vast scale sorrow up-and-comer datasets to uncover some fundamental online practice errors between discouraged clients and non-discouraged clients via web-based networking media. Since online practices can't be disregarded in current life, we anticipate that our discoveries should give more points of view and bits of knowledge to explore in software engineering and brain science.

Judy Hanwen Shen et al. [10] proposed past examinations concerning recognizing psychological instabilities through online networking have predominately centred on distinguishing melancholy through Twitter corpora (De Choudhury et al., 2013; Resnik et al., 2015; Pedersen, 2015). In this paper, we study uneasiness issues through close-to-home stories from the mainstream internet-based life site, Reddit. We construct a significant informational collection of run-of-the-mill and nervousness-related posts. We apply N-gram language displaying, vector embedding, point investigation, and passionate standards to produce includes that precisely characterize presents related to parallel degrees of uneasiness. We accomplish a precision of 91% with vectors pace word embedding and an exactness of 98% when joined with dictionary-based highlights. The LIWC 2015 lexicon adequately includes tension-related word use to effectively group Anxiety and Control Reddit posts. Be that as it may, by joining LIWC highlights with N-gram probabilities or unaided component age strategies (i.e., vector space inserting's and LDA Topic demonstrating), we can hoist the grouping exactness to 98%. Also, we discover connections between tension and explicit LDA subjects, for example, school and liquor (and medication) utilization (see Table 2). This could be a powerful strategy for distinguishing issues that individuals with tension or other psychological sicknesses talk about on the web.

Hayda Almeida et al. [4] expressed that this paper shows the frameworks created by the UQAM group for the CLEF in Pilot Task 2017. The objective was to foresee as right on time as conceivable the danger of emotional wellness issues from client-produced content in online networking. A few methodologies dependent on administered learning and data recovery techniques were utilized to assess the threat of melancholy for a client, given the substance of its posts on Reddit. The tests demonstrate that joining data recovery and AI methodologies gives the best outcomes among the five frameworks assessed. This report depicts the early hazard forecast frameworks

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

submitted to the CLEF eRisk 2017 pilot task. The framework that performed best depends on a multipronged approach, which consolidates expectations from SL and IR-based frameworks. SL-based frameworks utilized four significant element types and three order calculations, LMT, gathering SMO and outfit RF. IR-based frameworks use two records, and clients are positioned by a likeness score dependent on the BM25 positioning calculation. The expectations obtained from SL and IR-based frameworks converge by a choice calculation. The outcomes show that consolidating SL and IR methodologies beats the results obtained by each approach applied independently. In future work during our test stage, we have performed basic tests to assess the utilization of three different techniques: (1) basic standard-based arrangement utilizing a conclusion examination library, (2) profound learning-based order utilizing a Recurrent Neural Network (RNN), and (3) subject extraction utilizing Latent Dirichlet Allocation. Improving the framework execution will include further examination of these methodologies, just as an upgrade of the IR-based assets of the framework.

Long Ma et al. [3] proposed that a precise sorrow conclusion is a complex long-haul inquiry about the issue. The present discussion situated despondency conclusion between a medicinal specialist and an individual isn't exact because of the number of known indications. To find more presentations, our exploration work centres around removing elements identified with sorrow from web-based life, for example, informal organizations and websites. There are two significant points of interest in applying content mining instruments to new sorrow indications extraction. Right off the bat, individuals share their sentiments and information on Social Media. Besides, web-based life produces a massive volume of information that can be utilized for research reasons.

Kimberly McManus et al. [11] recommended that Individuals who experience the ill effects of schizophrenia include 1% of the United States populace and are multiple times more bound to pass on suicide than the general US populace. Recognizable proof of in-danger people with schizophrenia is testing when they don't seek treatment. Small-scale blogging enables clients to impart their contemplations and feelings to the world in short content. In this work, we utilized the vast corpus of Twitter pots and AI strategies to distinguish people with schizophrenia. Utilizing highlights from tweets, for example, Emoji use, posting time of day, and word reference terms, we prepared, manufactured and approved a few AI models. Our help vector machine model accomplished the best execution with 92% accuracy and 71% review on the held-out test set.

Moreover, we assembled a web application that powerfully shows outline statics between associates. This empowers efforts to undiscovered people, improved doctor analyses, and DE stigmatization of schizophrenia. This epic amalgamation of nostalgic examination strategies with enormous-scale Twitter information enabled us to distinguish beforehand described schizophrenic small-scale blogging propensities and precisely arrange Twitter clients with schizophrenia and control clients. Our investigation incorporated an associate of 96 twitter clients with schizophrenia and 200 age-coordinated controls. We utilized an SVM to isolate the associates dependent on their Twitter utilization designs with 92& exactness. At last, we made an intelligent information perception instrument to look at the outcomes, which are generalizable to

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

different undertakings. This work will have an extraordinary effect on the distinguishing proof of people with schizophrenia [12].

Amir Hossein Yazdavar et al. [13] suggested that with the ascent of web-based life, many individuals are routinely communicating their dispositions, sentiments, and daily battles with psychological wellness issues via web-based networking media stages like Twitter. Dissimilar to conventional observational companion studies directed through polls and self-detailed overviews, we investigate the solid recognition of clinical wretchedness from tweets got subtly. We showed the effect of web-based social networking on extraction and auspicious observing of gloom side effects. We built up a measurable model utilizing a half-and-half approach that consolidates a dictionary-based procedure with a semi-managed subject demonstrating a strategy to extricate per client point conveyance (clinical, symptomatic of despair) and per theme word dissemination (side effect markers) by printed investigation of tweets over various time windows. Our methodology supplements the present survey-driven indicative devices by gathering sadness side effects constantly and subtly. Our test results uncover critical contrasts in the subject inclinations and the word use example of one proclaimed discouraged gathering from arbitrary clients in our dataset, demonstrating the competency of our model for this assignment. Our model yields promising outcomes with an exactness of 68% and an accuracy of 72% for catching wretchedness manifestations per client over a period interim which is aggressive with a wholly managed methodology. In future, we intend to apply our way of dealing with different information sources, for example, longitudinal electronic well-being record (EHR) frameworks and personal protection repayment and case information, to build up a substantial "huge information" stage for identifying burdensome clinical conduct at the network level.

Victor Leiva et al. [14] expressed that, this paper is a starter endeavour to limit estimates that punish the postponement in distinguishing positive cases. Our tests underline the significance of a comprehensive supposition investigation and a mix of learning calculations to recognize early indications of gloom. His paper has explored how to all the more likely distinguish early danger of despair in online life by upgrading time-mindful order measures: ERDE5 and ERDE50. We have applied various learning calculations and mixes of them. Different procedures have been examined for dimensionality decrease and content extremity. We have given additional proof of the advantage of applying genetic analyses and content extremity (16, 7% improvement in the standard). Future work will focus on using progressively exact content conclusion grouping to show signs of improvement portrayal of the information highlights.

Farig Sadeque et al. [15] proposed The 2017 CLEF eRisk pilot assignment centres around consequently distinguishing despondency as right on time as conceivable from clients' presents on Reddit. This paper presents the procedures utilized for the University of Arizona group's interest in this early hazard discovery shared assignment. We used outer data past the little preparing set, including a previous despondency dictionary and ideas from the Unified Medical Language System as highlights. We used both successive (intermittent neural system) and nonconsecutive (bolster vector

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

machine) models for expectation. Our models perform reasonably on the test information, and the repetitive neural models perform superior to the non-successive help vector machines while utilizing similar capabilities. In this paper, we depicted the systems for the University of Arizona entries to the 2017 CLEF eRisk early hazard identification of melancholy pilot task. We utilized highlights dependent on a downturn dictionary and the Unified Medical Language System (UMLS). We executed consecutive and non-successive models and utilized gathering strategies to use the best of each model. We found that the group model works superior to the individual models, and sitting tight for more information before settling on a choice improves the customary exhibition estimates like exactness, review and F1. Regardless of whether it is adequate to hold up a good measure of time to have better execution is as yet an open inquiry – and we might want to chip away at that. We might want to set up a time allotment that can be esteemed worthy before settling on a choice with the goal that the tradeoff between the accuracy and speed of the option is limited.

Ahmed Husseini Orabi et al. [15] suggested that mental ailment location in internet-based life can be viewed as a perplexing errand, primarily because of the convoluted idea of mental issues. As of late, this examination zone has begun to develop with the constant increment in the ubiquity of online life stages that have become a necessary piece of individuals' lives. This cozy connection between web-based life stages and clients has made them mirror the clients' close-to-home life on numerous levels. In such a situation, analysts are given much data concerning one's life. Notwithstanding the multifaceted nature of distinguishing psychological sicknesses through web-based social networking stages, receiving regulated AI methodologies, for example, profound neural systems, have not been generally acknowledged because of the challenges in acquiring adequate measures of commented-on preparing information. Because of these reasons, we attempt to distinguish the best profound neural system design among a couple of chosen structures that were effectively utilized in everyday language preparation errands. The picked methods are used to recognize clients with indications of psychological maladjustments (gloom for our situation) given restricted unstructured content information extricated from the Twitter internet-based life stage. All in all, we displayed a novel way to deal with streamlined word installation for characterization errands.

3. PROPOSED METHODOLOGY

The research productivity proposed a technique and solved an existing problem. People have worked on depression detection in social media users of specific cultures, areas, races and countries. All research is done using Text-based data, but no research is done using emojis. So I am providing a solution by using text with emojis. Emojis play a vital role and express feelings, expressions, thoughts, vision and emotions better than textual data. Sometimes people post statuses with just emojis without any text representation. So it helps in understanding situations. Our primary purpose and objective are to identify depression-affected users early to diagnose and control it. We will provide general solutions for all kinds, i.e. race, culture, and area of people. To achieve this motivation, we use intelligent methods of machine learning. We have

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

studied six techniques to evaluate the twitter dataset in the literature review. After reviewing the results of these techniques, we observed variations in results. First of all, people use the individual classifier.

Then we faced an issue of how to improve its performance, so to improve performance, we had to use more than one classifier as we had already studied individual classifiers and seen their results. There are two kinds of classifiers. Homogenous and heterogeneous. In a similar homogenous type of classifiers are used. It has two different types Bagging and Boosting. Most people use individual classifiers for data analysis. This way, we have to use multiple classifiers and compare outcomes to find better results. We have used and seen individual classifiers' results, so we used the bagging classifier. It's best to improve results because it takes cumulative effect by combining results of all the exact nature of Bagging Classifiers. It is a combiner and is applied to developing the same kind of classifiers. Different heterogeneous classifiers are used, and stacking is applied. I am just introducing and giving the concept of both kinds of classifiers. But according to my research domain Homogenous kind of classifier, I moved forward keeping the features and advantages of bagging classifier in view. Our Analysis and experiments found that Bagging Classifier is better and yields promising results compared to the existing classifier.

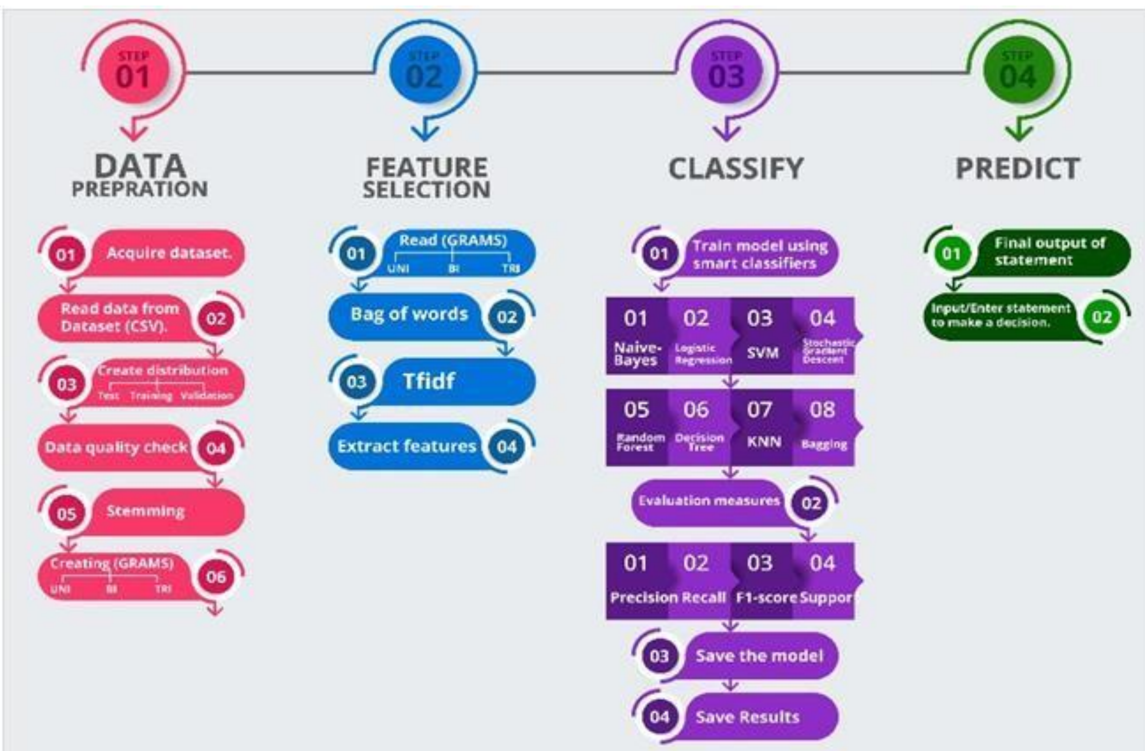

Figure 1: Proposed methodology

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

3.1 Data preparation

It is the first and essential step to start because without data, and we are blind to go ahead. It plays a vital role in the analysis. We selected Twitter as a data source because it's more authentic, reliable, and readily available. By keeping in previous focus research, we found almost 85% of research using the Twitter dataset. To get data, we download it through Twitter API. When we downloaded the data set successfully, it was almost in CSV format. I also need it in this format, so I don't have to convert or change it. The next step is feature extraction and reduction to extract useful information from raw data. Then divide this data into three sets test, training and validation. Apply preprocessor to process data feature analysis. To create a distribution, we split 80% data for training and 20% for testing. We also add some quality checks to improve the information. Data preprocessing is converting raw data into an understandable format and valuable information.

Stemming is a component of linguistic study in morphology and information extraction and retrieval using artificial intelligence (AI). Since other word forms linked with a subject may need to be searched to acquire better results, stemming and AI information extract useful statistics from vast sources like big data or the Internet. Finding, probing, and recovering more word forms produces different outcomes. When a word form is predictable, it can deliver search results that might otherwise be squandered. Due to the additional information that may be retrieved, stemming is essential for search queries and information recovery. A recently identified word could provide fresh research opportunities. The best results are frequently obtained using the lemma, the word's basic morphological form. The lemma is determined through stemming, which a person or an algorithm can do that an AI system may employ. Any inflected form encountered can be stemmed back to its basic form. Generating grammes is the sixth phase. The three types of grammes are UNI, BI, and Tri Gram.

3.2 Feature selection

In feature selection, we first read UNI, bi and tri grams categories. It plays an essential role in creating a bag of words. It is the processing for choosing features for analysis. The feature means words, keywords or sentiments. In this process, we create a list of words to make a master list which is like a validation check to pass through for each message and tweet. The selection of features is an essential part of learning the machine. Selection of features refers to reducing processing and analysing inputs or finding the most meaningful inputs. A similar term, software engineering (or extraction of features), refers to extracting valuable data or functionality from existing data. Information almost always includes more information than is required to create the model or the wrong type of information. For several reasons, selecting features is critical to building a good model. One is that selection of parts implies some degree of reduction in cardinality to impose a cutoff on the number of attributes that can be considered when constructing a model. Not only does the selection of features improve the model's quality, but it also makes the modelling process more efficient. If you use unnecessary columns while creating a model, the training process requires more CPU and memory, and the completed model requires more storage space. Even if resources

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

weren't a concern, you still want to pick a function and define it. A bag of words is a collection of depression-related or positive comments.

After preparation and preprocessing of data to get data in an understandable form. We label data by assigning IDs, 1 for positive, -1 for negative and 0 for neutral messages. This labelling technique is used to create the bag of words; these labels define the scope of the tweet, whether it lies in a depressed or happy word cloud. A load of comments or word cloud is the same, showing a clear image of the data set and the ratio of depression and happiness in the diagram. N-grams also play an essential role in creating a bag of words. Now I will explain Tf-idf. Tf-idf stands for term frequency Inverse document frequency, and the Tf-idf weight is often used in information retrieval and text mining. Here we train the model and then give input in tweet to test whether it is positive or negative. Positive means this tweet holder is a depressed person, and negative is vice versa.

TF is Term frequency means the number of iterations a word appears in a document divided by the total number of words in the document. IDF mean inverse data frequency is the log of the number of documents divided by some documents. It concludes the weight of rare words across all documents in quantity. TF-IDF (Term Frequency-Inverse Document Frequency) is a statistical measure that assesses the relevance of a comment to a document in a document collection. This is done by multiplying two metrics: how many times a word in a document appears and the word's inverse document frequency across a set of documents. The last step in this phase is feature selection and reduction. There are two categories of features: feature selection and feature reduction. Feature selection is self-explanatory; it means which features are best and concern to select. The other term, feature reduction, means which features to reduce, remove or skip for analysis.

3.3 Classify

A classifier uses training data to understand how the class relates to given input variables. When the classifier is correctly trained, an unknown email can be identified. Classification belongs to the supervised learning category, where the targets provide input data. The type belongs to the supervised learning group, where the goals provide input data. There are many applications in different domains. There are a lot of applications in classification for other parts. So we will train a model using intelligent classifiers. There are a lot of classification algorithms existing now, but it is not possible to determine which one is better than the other. It depends on the usage and quality of the data set available. For Instance, if the classes are linearly discrete, linear classifiers like Logistic regression, Fisher’s linear discriminant can overtake sophisticated models and vice versa.

3.4 Results and Discussions

This chapter presents different results achieved using various machine learning techniques to predict how we can use online data to evaluate depressed people using the Twitter dataset. As we have already discussed, the primary purpose of this thesis is to detect depressed people initially. We have used different intelligent methods and classifiers to get better and more promising results. Bagging is the best classifier I have ever experienced to yield maximum accuracy percentages, Precision, Recall and F-Score. I have applied eleven classifiers and shared their results in the figures and tabular form below. We tried eleven classifiers with six data sets one by one and got different results. Our main objective was to maximize results.

Table 2: Data sets

Sr. No	Name	Class/Section	Source	Published
1	Sentimental Analysis with Tweets	Twitter	[21]	2017
2	An advanced Twitter scraping	Twitter	[22]	2017
3	Detecting depression in Tweets	Twitter	[23]	2018
4	Depression Sentiment Analysis with Twitter-Data	Twitter	[24]	2017
5	Early risk perdition on internet	CLEF	[25]	2017
6	Classification of Depression on Social Media Using Text Mining	Twitter	[26]	2018

3.5 Data Collection

We collected different data sets and projects from GitHub and Kaggle Websites, which have worked in this domain. Then executed those data sets and found their findings. After that observed their weakness and made a comparison of the results. Table 1 shows the collection of datasets, project name, sources and year of publication. To achieve the objectives of this thesis. The first step is choosing social media with more accurate and authentic data. As we all know, there are some fake accounts, or some people have multiple accounts for different purposes. There is a vast quantity of data available on social media, which helps detect depressed people, as each dataset has its format and structure. There are many social media sites and apps to collect data using the internet worldwide. According to my research, I found progressively exact and bona fide information on Twitter for Depression examination. So In the very first step, download Twitter data using Twitter API, then test this data set. When a data set is downloaded, pass this dataset to the preprocessor to make it understandable and remove noisy data like extra spaces and symbols. Data preprocessing is self-explanatory terminology, a mining technique to transform raw data into a legible/readable format. Accurate world data is inconsistent primarily, incomplete, containing extra spaces symbols, unnecessary punctuation and symbols which sometimes make no sense containing many errors and lacking specific trends and behaviours. It is a well-known technique to resolve such kinds of issues. It prepares such types of data for further processing. Data pass through some steps during preprocessing. Those steps are data cleaning, data integration, data transformation, data reduction and data discretization.

After successful preprocessing, we do despair sentiment evaluation with the aid of the going for walks script function, which will go over the output document while retrieving

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

the tweet matching each sentiment's identification. We employ these facts as input for our classifiers. When everything is finished, the console must include a list of each classifier's AUC (Area under the Curve). So I collected many twitter informational indexes and finished six to do various things with. After gathering facts, I applied a preprocessor to evacuate boisterous facts like different snapshots, emojis, and areas. Information Pre-making ready is a strategy used to evacuate, clean and convert crude and unpleasant information into essential informational series. Dissimilar to the highlight extraction approach, determination strategies accumulate any other summed-up list of abilities from the first set. This subset of highlights chosen offers first-class execution due to some goal paintings and memorable criteria. In the wake of having treasured and essential facts, we do a wistful investigation by applying multiple classifiers, but we get maximum evaluated consequences from the packing classifier.

4. Results and Discussion

The above table shows a collection of data sets used from 2017 to 2018. Most data sets are of Twitter because I have already described that Twitter is the most used social media for sentimental analysis. It is a reliable and authentic source of data. Almost research is done using Twitter datasets.

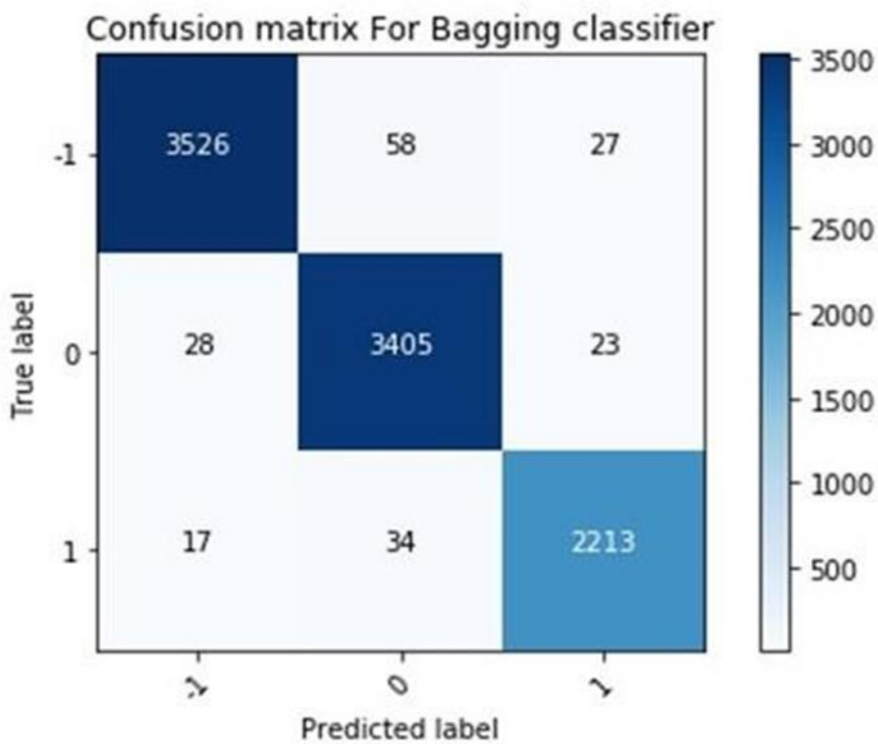

Figure 2: Confusion Matrix for Bagging Classifier

A confusion matrix, also known as an error matrix, is used in machine learning, especially in the challenge of statistical classification. A table known as a confusion matrix is frequently used to illustrate how a classification model, also known as a

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

"classifier," performed on a set of test data for which the actual values were known. It enables the viewing of an algorithm's performance. It allows simple identification of class labelling confusion, such as when one class is frequently mislabeled as the other. The confusion matrix is used to calculate the majority of performance metrics. A confusion matrix gathers the anticipated results of a classification issue. The number of accurate and incorrect predictions for each class is expressed using count values. This is the secret of the confusion matrix. The confusion matrix shows how your classification model makes projections while confused. More importantly than simply the faults a classifier is making, it provides us with information on the different types of errors that are being made.

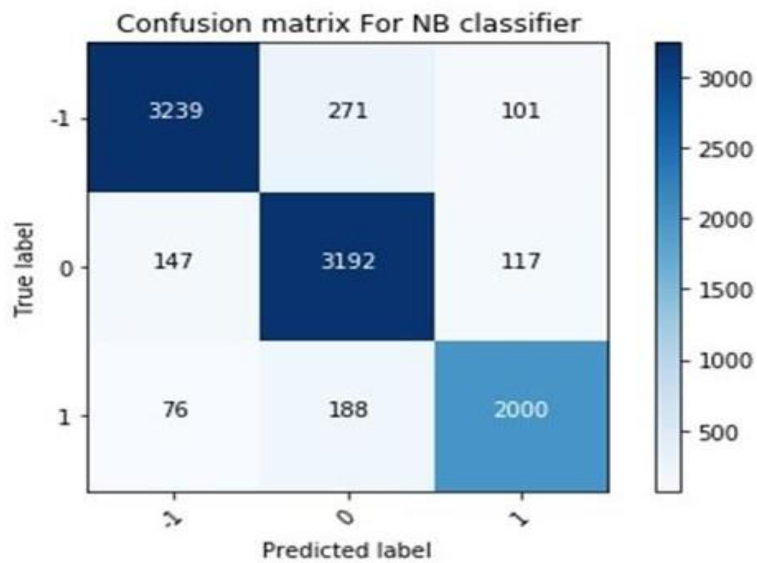

Figure 3: Confusion Matrix for Naïve Bayes Classifier

A classifier's output quality on a data set is assessed using a confusion matrix. The number of points for which the predicted label matches the actual label is shown by the diagonal elements, whilst the off-diagonal elements represent the incorrect labels assigned by the classifier. The confusion matrix's diagonal values should be high, suggesting numerous accurate predictions. Above, both figures show the Confusion matrix for Bagging and Naïve Bayes classifiers. The following are terms used for analysis.

- Positive (P): Observation is positive.
- Negative (N): Observation is negative.
- True Positive (TP): Observation is positive and is predicted to be positive.
- False Negative (FN): Observation is positive but is predicted negative.
- True Negative (TN): Observation is negative and is predicted to be negative.
- False Positive (FP): Observation is negative, but is predicted positive.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

Figure 4: Happiness Bag of words (BOG)

The above figure shows a cluster of positive words created from most Twitter datasets as we have downloaded the Twitter dataset. After preprocessing, we achieved meaningful and accurate data in an understandable form. After successful preprocessing, we split data into training and testing sets. We create a table containing three columns. The first column contains an index of every tweet, the second has a tweet message, and in the third column, we assign a label to the tweet, either positive, negative or neutral. Now we will use this table to create word cloud analysis and check if the tweet label is zero, then find the positive word in a tweet and set it in the list. In this, we get a list of positive comments, then call a figure and show it by customizing its properties. We may customize the size of this figure in both the x and y frame axis and the colour of words etc.

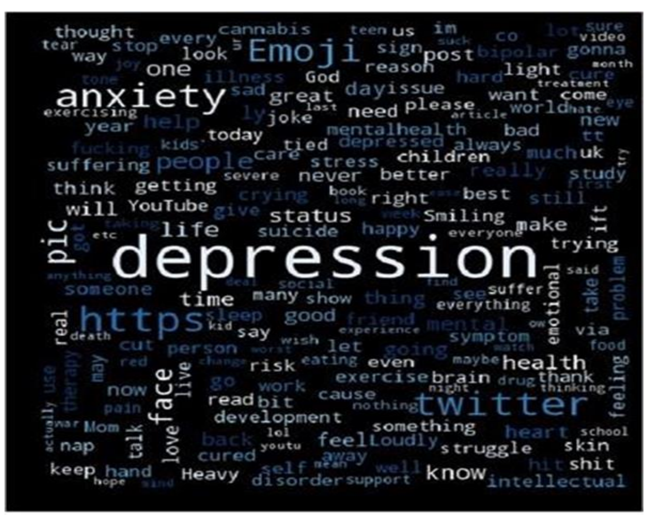

Figure 5: Depression Bag of words (BOG)

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

The bag of words shows a cluster of depressive words created from a bulk of Twitter datasets as we have downloaded the Twitter dataset. After preprocessing, we achieved meaningful and accurate data in an understandable form. After successful preprocessing, we split data into training and testing sets. We create a table containing three columns.

The first column contains an index of every tweet, the second has a tweet message, and in the third column, we assign a label to the tweet, either positive, negative or neutral. Now we will use this table to create word cloud analysis and check if the tweet label is one, then find a positive word in a tweet and set it in the list. In this, we get a list of positive comments, then call the figure and show it by customizing its properties.

We may customize the size of this figure in both the x and y frame axis and the colour of words etc. All procedure for drawing both figures is the same except for the label, as we are distinguishing happy and depressed bag of comments from the title of tweets. So labelling plays the leading role in identifying whether either tweet lies in a happy or depressive cloud.

Table 3: Results obtained from different classifiers

Classifier	Accuracy	Precision	Recall	F-score
Naive Bayes	82.58%	76.15%	50.15%	65.78%
Decision Tree	89.33%	63.53%	83.14%	75.33%
Support Vector Machine	50.00%	62.25%	87.74%	77.41%
Kneighbors Classifier	85.58%	79.49%	87.74%	82.43%
Random Forest	68.44%	79.42%	76.82%	84.32%
Logistic Regression	95.90 %	81.25%	73.84%	86.30%
Bagging Model	98.33%	90.39%	96.45%	92.15%

Table 3 shows a list of Naïve Bayes, Decision trees, Support vector machines, Kneighbors, Random forests, Logistic regression and bagging classifiers. All columns against these classifiers show results in terms of Accuracy, Precision, Recall and F-Score. A comparison of the above table results shows that the bagging classifier ranks first in terms of all factors. All classifiers are machine-learning intelligent techniques and of homogenous nature.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

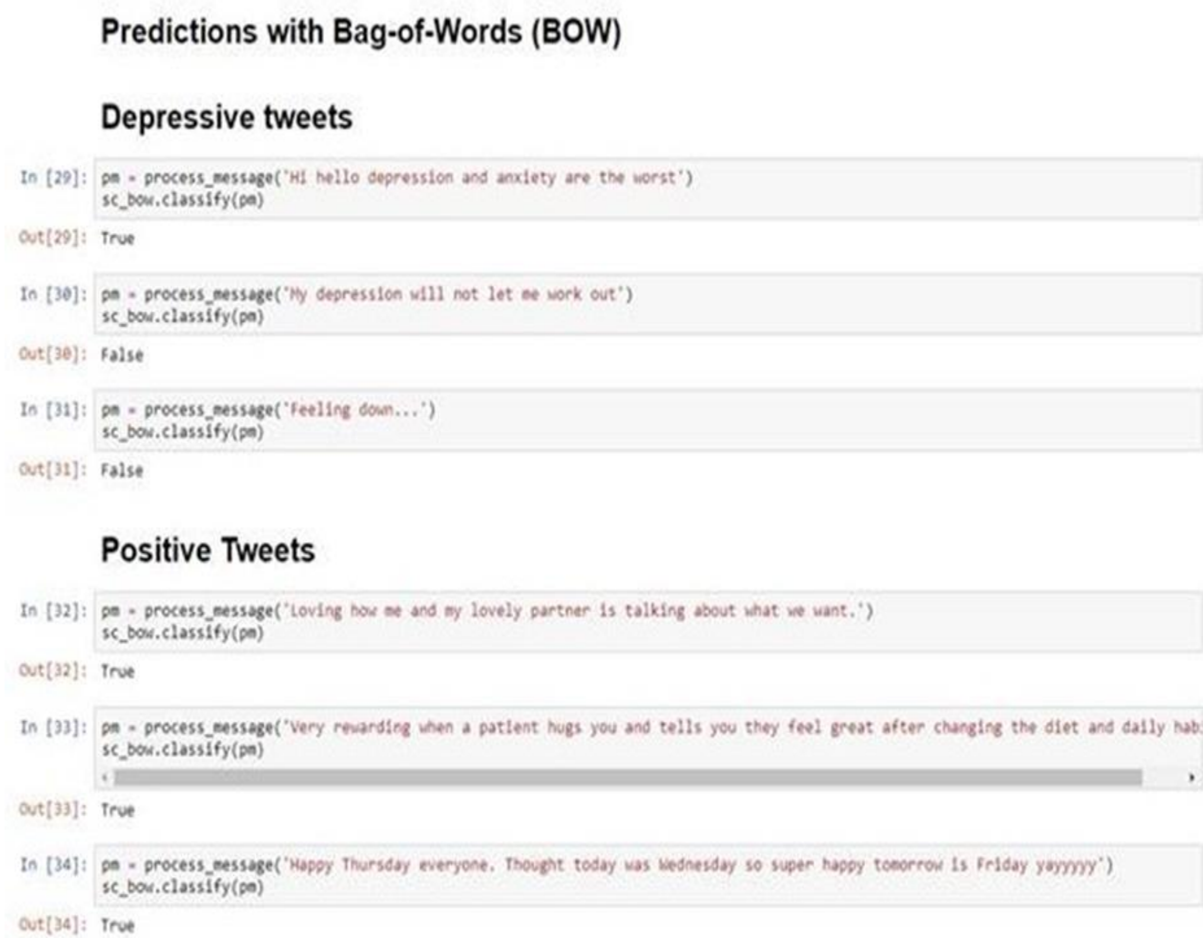

Figure 6: Prediction with BOW Depressive and Positive Tweets

Figure 6 presents an analysis of tweets, and firstly I have created a bag of words cloud for depressive and happy tweets. Now my model is trained for research, so I input a tweet and test result to check whether it's negative or positive. It's called prediction with a Bag of words with depressive and positive tweets. First, we have to train maximum data to train the model and then test it, and our test results depend on trained data sets. So to get better results, understanding and accuracy of data sets and situations. Because as more data is taught, it covers maximum possibilities and scenarios. Both figures show positive and negative tweets. A depressive tweet genuinely means it is depressive and negative means depression free. In positive tweets, it is directly proportional true means positive and false means negative.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

Figure 7: Prediction with TF-IDF Depressive Tweets

Figure 7 shows the results of the prediction with Tf-Idf depression tweets. Tf-idf stands for term frequency Inverse document frequency, and the Tf-idf weight is often used in information retrieval and text mining. Here we train the model and then give input in tweet to test whether it is positive or negative. Positive mean this tweet holder is a depressed person, and negative is vice versa.

Figure 8: Prediction with TF-IDF Positive Tweets

Figure 8 also describes prediction with Tf-Idf depression tweets. Tf-idf means term frequency. Inverse document frequency and Tf-idf weight are often used in information retrieval and text mining. It is also positive and direct relation that demonstrates that true means its depression free and vice versa.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 01-2023
DOI 10.17605/OSF.IO/EVXB7

5. CONCLUSION

In this research, we have measured research progress under machine learning techniques and intelligent methods using social media data to detect depressed people. I have looked into different social media to find better data sets for experiments. Previous research showed more research on Twitter datasets. So my personal experience also agrees that the Twitter dataset is the best, most accurate and most reliable source. After finalizing data sets and considering our goal, I implemented intelligent machine learning techniques and mentioned results in figures and tables. I found different results using available data sets of all kinds of users, i.e. this data and solution are generalized involving users of other races, cultures, cities and countries. According to our sole purpose, we have achieved maximum accuracy by using the Homogenous nature of the classifier known as bagging.

Additionally, we suggested the Bag of Words method for analysing this dataset, using dimensional feature space as our input vector. In the end, we used these distinguishing characteristics to construct, evaluate, and contrast several statistical classifiers that may be able to predict the possibility of depression in a person at an early stage. We found the Bag of Words approach to be a proper feature set and bigrams to offer no discernible benefit over a unigram-based approach.

In our upcoming study, we seek to comprehend how spatiotemporal behaviour may contribute to the growth of Major Depressive Disorder. Evaluation, conclusion, and inference of daily fluctuations in depression may be a valuable technique for spotting depression before moderate beginnings, increasing its potential to save lives. It is crucial to discover methods that might be applied in a medical setting to recognize clinical depression from social media users' behaviour to benefit the general population.

References

❖ ADDIN ZOTERO_BIBL {"uncited":[],"omitted":[],"custom":[]} CSL_BIBLIOGRAPHY [1] M. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, “Depression detection from social network data using machine learning techniques,” Health Inf. Sci. Syst., vol. 6, no. 1, pp. 1–12, 2018.

❖ M. Nadeem, “Identifying depression on Twitter,” ArXiv Prepr. ArXiv160707384, 2016.

❖ L. Ma, Z. Wang, and Y. Zhang, “Extracting depression symptoms from social networks and web blogs via text mining,” presented at the International Symposium on Bioinformatics Research and Applications, 2017, pp. 325–330.

❖ H. Almeida, A. Briand, and M.-J. Meurs, “Detecting Early Risk of Depression from Social Media User-generated Content.,” presented at the CLEF (Working Notes), 2017.

❖ H. Orabi, P. Buddhitha, M. H. Orabi, and D. Inkpen, “Deep learning for depression detection of twitter users,” presented at the Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, 2018, pp. 88–97.

❖ M. J. Vioules, B. Moulahi, J. Azé, and S. Bringay, “Detection of suicide-related posts in Twitter data streams,” IBM J. Res. Dev., vol. 62, no. 1, pp. 7–1, 2018.

❖ F. Sadeque, D. Xu, and S. Bethard, “Measuring the latency of depression detection in social media,” presented at the Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 495–503.

❖ T. Shen et al., “Cross-domain depression detection via harvesting social media,” 2018.

❖ G. Shen et al., “Depression detection via harvesting social media: A multimodal dictionary learning solution.,” presented at the IJCAI, 2017, pp. 3838–3844.

❖ J. H. Shen and F. Rudzicz, “Detecting anxiety through reddit,” presented at the Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology—From Linguistic Signal to Clinical Reality, 2017, pp. 58–65.

❖ K. McManus, E. K. Mallory, R. L. Goldfeder, W. A. Haynes, and J. D. Tatum, “Mining Twitter data to improve detection of schizophrenia,” AMIA Summits Transl. Sci. Proc., vol. 2015, p. 122, 2015.

❖ Z. Jamil, “Monitoring tweets for depression to detect at-risk users,” 2017.

❖ H. Yazdavar et al., “Semi-supervised approach to monitoring clinical depressive symptoms in social media,” presented at the Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, 2017, pp. 1191–1198.

❖ V. Leiva and A. Freire, “Towards suicide prevention: early detection of depression on social media,” presented at the International Conference on Internet Science, 2017, pp. 428–436.

❖ F. Sadeque, D. Xu, and S. Bethard, “UArizona at the CLEF eRisk 2017 pilot task: linear and recurrent models for early depression detection,” presented at the CEUR workshop proceedings, 2017, vol. 1866.

❖ K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot, and Y. Kaewpitakkun, “Facebook social media for depression detection in the Thai community,” presented at the 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2018, pp. 1–6.

❖ Y. Wang, Z. Wang, C. Li, Y. Zhang, and H. Wang, “Online social network individual depression detection using a multitask heterogenous modality fusion approach,” Inf. Sci., vol. 609, pp. 727–749, 2022.

❖ Y. Hswen, J. A. Naslund, J. S. Brownstein, and J. B. Hawkins, “Online communication about depression and anxiety among twitter users with schizophrenia: preliminary findings to inform a digital phenotype using social media,” Psychiatr. Q., vol. 89, no. 3, pp. 569–580, 2018.

❖ Martínez, R. Benítez, H. Estrada, and Y. Hernández, “Predictive model for detection of depression based on uncertainty analysis methods,” Multidiscip. Digit. Publ. Inst. Proc., vol. 2, no. 19, p. 551, 2018.

❖ Shensa, J. E. Sidani, M. A. Dew, C. G. Escobar-Viera, and B. A. Primack, “Social media use and depression and anxiety symptoms: A cluster analysis,” Am. J. Health Behav., vol. 42, no. 2, pp. 116–128, 2018.

❖ M. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, “Depression detection from social network data using machine learning techniques,” Health Inf. Sci. Syst., vol. 6, no. 1, pp. 1–12, 2018.

❖ Y. Hswen, J. A. Naslund, J. S. Brownstein, and J. B. Hawkins, “Online communication about depression and anxiety among twitter users with schizophrenia: preliminary findings to inform a digital phenotype using social media,” Psychiatr. Q., vol. 89, no. 3, pp. 569–580, 2018.

❖ Iqbal, M. J., Bajwa, U. I., Gilanie, G., Iftikhar, M. A., & Anwar, M. W. (2022). Automatic brain tumor segmentation from magnetic resonance images using superpixel-based approach. Multimedia Tools and Applications, 1-19.

❖ Muhammad Javaid Iqbal Et Al, .. (2022, December 10). Blindness Detection Using Machine Learning Approaches. https://doi.org/10.17605/OSF.IO/T3XDU