Investigation of Social Media on Depression Using Machine Learning

Dr.R. Vijayarajeswari[1], Praveenkumar.M[2], Mano.K[3], Dhanapal.P[4]
[1] *Associate Professor, Department of Computer Science and Engineering, Mahendra Engineering College, Tamilnadu, India*
[2,3,4] *Ug Students, Department of Computer Science and Engineering, Mahendra Engineering College, Tamilnadu, India*

Abstract **- The explosive growth in popularity of social networking leads to the problematic usage. An increasing number of social network mental depression (SNMDs), such as Cyber-Relationship Addiction, Information Overload, and Net Compulsion, have been recently noted. Symptoms of these mental depression are usually observed passively today, resulting in delayed clinical intervention. In this Project, we argue that mining online social behavior provides an opportunity to actively identify SNMDs at an early stage. It is challenging to detect SNMDs because the mental status cannot be directly observed from online social activity logs. Our approach, new and innovative to the practice of SNMD detection, does not rely on self-revealing of those mental factors via questionnaires in Psychology. Instead, we propose a machine learning framework, namely, Social Network Mental Depression Detection (SNMDD) that exploits features extracted from social network data to accurately identify potential cases of SNMDs. With rapid technological advancement, many have questioned the benefits and side effects of social media on a user's psychological health. A study was carried out in India to determine the relationship of social media usage and depression in Asian individuals of different genders and age groups. Our findings show that the higher the usage of social media, the higher the risk of depression, with teenage girls being subjected to the highest risk. We also exploit multi-source learning in SNMDD and propose a new SNMD-based Tensor Model (STM) to improve the accuracy. To increase the scalability of STM, we further improve the efficiency with performance guarantee. Our framework is evaluated via a user study with 3126 online social network users. We conduct a feature analysis, and also apply SNMDD on large-scale datasets and analyze the characteristics of the three SNMD types. The results manifest that SNMDD is promising for identifying online social network users with potential SNMDs.**

Index Terms **- Deep neural network, context, depression risk, mental health, multiple regression, healthcare, deep learning, context information.**

1.INTRODUCTION

1.1 DEEP NEURAL NETWORK

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.

Deep learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, machine vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance. Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems. ANNs have various differences from biological brains. Specifically, neural networks tend to be static and symbolic, while the biological brain of most living organisms is dynamic (plastic) and analog. The adjective "deep" in deep learning comes from the use of multiple layers in the network. Early work showed that a linear perceptron cannot be a universal classifier, and then that a network with a nonpolynomial activation function with one hidden layer of unbounded width can on the other hand so be. Deep learning is a modern variation which is

concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically informed connectionist models, for the sake of efficiency, trainability and understandability, whence the "structured" part.

1.2 DEPRESSION RISK

Depression is a mood Depression that causes a persistent feeling of sadness and loss of interest. Also called major depressive Depression or clinical depression, it affects how you feel, think and behave and can lead to a variety of emotional and physical problems. You may have trouble doing normal day-to-day activities, and sometimes you may feel as if life isn't worth living. More than just a bout of the blues, depression isn't a weakness and you can't simply "snap out" of it. Depression may require long-term treatment. Depression may occur only once during your life; people typically have multiple episodes. During these episodes, symptoms occur most of the day, nearly every day and may include. For many people with depression, symptoms usually are severe enough to cause noticeable problems in day-to-day activities, such as work, school, social activities or relationships with others. Depression is not a normal part of growing older, and it should never be taken lightly. Unfortunately, depression often goes undiagnosed and untreated in older adults, and they may feel reluctant to seek help. Depression is a serious Depression that can take a terrible toll on you and your family. Depression often gets worse if it isn't treated, resulting in emotional, behavioral and health problems that affect every area of your life.

1.3 MENTAL HEALTH

Mental health is the level of psychological well-being or an absence of mental illness. It is the state of someone who is "functioning at a satisfactory level of emotional and behavioral adjustment". From the perspectives of positive psychology or of holism, mental health may include an individual's ability to enjoy life and to create a balance between life activities and efforts to achiev psychological resilience. According to the World Health Organization (WHO), mental health includes "subjective well-being, perceived self-efficacy, autonomy, competence, intergenerational dependence, and self-actualization of one's intellectual and emotional potential, among others". The WHO further states that the well-being of an individual is encompassed in the realization of their abilities, coping with normal stresses of life, productive work, and contribution to their community. Cultural differences, subjective assessments, and competing professional theories all affect how one defines "mental health"

1.4 MULTIPLE REGRESSION

Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The variable we want to predict is called the dependent variable (or sometimes, the outcome, target or criterion variable). The variables we are using to predict the value of the dependent variable are called the independent variables (or sometimes, the predictor, explanatory or regressor variables). For example, you could use multiple regression to understand whether exam performance can be predicted based on revision time, test anxiety, lecture attendance and gender. Alternately, you could use multiple regression to understand whether daily cigarette consumption can be predicted based on smoking duration, age when started smoking, smoker type, income and gender. Multiple regression also allows you to determine the overall fit (variance explained) of the model and the relative contribution of each of the predictors to the total variance explained. For example, you might want to know how much of the variation in exam performance can be explained by revision time, test anxiety, lecture attendance and gender "as a whole", but also the "relative contribution" of each independent variable in explaining the variance. This "quick start" guide shows you how to carry out multiple regression using SPSS Statistics, as well as interpret and report the results from this test. However, before we introduce you to this procedure, you need to understand the different assumptions that your data must meet in order for multiple regression to give you a valid result. We discuss these assumptions next

1.5 DEEP LEARNING CONTEXT INFORMATION

The introduction of deep neural networks (DNNs) has advanced the performance of automatic speech recognition (ASR) tremendously. On a wide range of ASR tasks, DNN models show superior performance than the traditional Gaussian mixture models (GMMs). Although making significant advances, DNN models still suffer from data scarcity, speaker mismatch and environment variability. This thesis resolves these challenges by fully exploiting DNNs' ability of integrating heterogeneous features under the same optimization objective. We propose to improve DNN models under these challenging conditions by incorporating context information into DNN training. On a new language, the amount of training data may become highly limited. This data scarcity causes degradation on the recognition accuracy of DNN models. A solution is to transfer knowledge from other languages to the low-resource condition. This thesis proposes a framework to build cross-language DNNs via languageuniversal feature extractors (LUFEs). Convolutional neural networks (CNNs) and deep maxout networks (DMNs) are employed to improve the quality of LUFEs, which enables the generation of invariant and sparse feature representations. This framework notably improves the recognition accuracy on a wide range of low-resource languages. The performance of DNNs degrades when the mismatch between acoustic models and testing speakers exists. A form of context information which encapsulates speaker characteristics is i-vectors. This thesis proposes a novel framework to perform feature-space speaker adaptive training (SAT) for DNN models. A key component of this approach is an adaptation network which takes i-vectors as inputs and projects DNN inputs into a normalized feature space. The DNN model fine-tuned in this new feature space rules out speaker variability and becomes more independent of specific speakers. This SAT method is applicable to different feature types and model architectures. The proposed adaptive training framework is further extended to incorporate distance- and video-related context information. The distance descriptors are extracted from deep learning models which are trained to distinguish distance types on the frame level. Distance adaptive training (DAT) using these descriptors captures speaker-microphone distance dynamically on the frame level. When performing ASR on video data, we naturally have access to both the speech and the video modality.

2. LITERATURE REVIEW

CONTEXT MINING BASED MENTAL HEALTH MODEL FOR LIFECARE PLATFORM - Author: Ji-Won Baek, Hoill Jung, Kyungyong Chung

With the emergence of the 4th industrial revolution, IT convergence engineering based artificial intelligence and intelligent system has constantly been researched in today's society. In particular, healthcare service based on IT-BT convergence helps to improve quality of people's life and provide user-oriented healthcare contents actively. Currently, the healthcare industry has gradually changed its healthcare paradigm from conventional healthcare to mental diseases care and tries to solve the social problem with depression, one of mental Depressions. This study makes use of users' profiles about depression and health weather index provided by Korea Meteorological Administration to classify and define semantic ontology-based context information, and to develop the context mining model for depression index service. The proposed context mining based mental health model uses personalized context information so that it is possible to provide personalized depression index service, rather than unified healthcare service

PREDICTION MODEL OF USER PHYSICAL ACTIVITY USING DATA CHARACTERISTICS-BASED LONG SHORT-TERM MEMORY RECURRENT NEURAL NETWORKS - Author: Joo-Chang Kim, Kyungyong Chung

Mobile healthcare services have attracted significant attention because of the emerging development and supply of diverse wearable devices. Smart watches and health bands are the most common type of mobile-based wearable devices and their market size is increasing considerably. However, simple value comparisons based on accumulated data have revealed certain problems, such as the standardized nature of health management and the lack of personalized health management service models. The convergence of information technology (IT) and biotechnology (BT) has shifted the medical paradigm from continuous health management and disease prevention to the development of a system that can be used to provide ground-based medical services regardless of the user's location.

GENERALIZING SEMI-SUPERVISED GENERATIVE DVERSARIAL NETWORKS TO REGRESSION USING FEATURE CONTRASTING - Author: Greg Olmschenk, Zhigang Zhu, Hao Tang

We generalize semi-supervised generative adversarial networks (GANs) from classification problems to regression problems. In the last few years, the importance of improving the training of neural networks using semi-supervised training has been demonstrated for classification problems. We present a novel loss function, called feature contrasting, resulting in a discriminator which can distinguish between fake and real data based on feature statistics. This method avoids potential biases and limitations of alternative approaches. The generalization of semi-supervised GANs to the regime of regression problems of opens their use to countless applications as well as providing an avenue for adeeper understanding of how GANs function. We first demonstrate the capabilities of semi-supervised regression GANs on a toy dataset which allows for a detailed understanding of how they operate in various circumstances.

PROPERTIES OF THE SIGN GRADIENT DESCENT ALGORITHMS - Author: Emmanuel Moulay, Vincent Lechappe, Franck Plestan

The aim of this article is to study the properties of the sign gradient descent algorithms involving the sign of the gradient instead of the gradient itself and _rest introduced in the RPROP algorithm. This article provides two results of convergence for local optimization, a _rst one for nominal systems without uncertainty and a second one for systems with uncertainties. New sign gradient descent algorithms including the dichotomy algorithm DICHO are applied on several examples to show their electiveness in terms of speed of convergence. As a novelty, the sign gradient descent algorithms can allow to converge in practice towards other minima than the closest minimum of the initial condition making these algorithms suitable for global optimization as a new met heuristic method.

MULTI-TASK LEAST SQUARES TWIN SUPPORT VECTOR MACHINE FOR CLASSIFICATION - Author: Benshan Mei, Yitian Xu

With the bloom of machine learning, pattern recognition plays an important role in many aspects. How- ever, traditional pattern recognition mainly focuses on single task learning (STL), and the multi-task learning (MTL) has largely been ignored. Compared to STL, MTL can improve the performance of learning methods through the shared information among all tasks. Inspired by the recently proposed directed multi-task twin support vector machine (DMTSVM) and the least squares twin support vector ma- chine (LSTWSVM), we put forward a novel multi-task least squares twin support vector machine (MTLS- TWSVM). Instead of two dual quadratic programming problems (QPPs) solved in DMTSVM, our algorithm only needs to deal with two smaller linear equations. This leads to simple solutions, and the calculation can be effectively accelerated. Thus, our proposed model can be applied to the large scale datasets. In addition, it can deal with linear inseparable samples by using kernel trick.

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

Internet Addiction Depression (IAD) is a type of behavior addiction with the patients addicted to the Internet, just like those addicting to drugs or alcohol. Many research works in Psychology, and Psychiatry have studied the important factors, possible consequences, and correlations of IAD. King investigate the problem of simulated gambling via digital and social media to analyze the correlation of different factors, e.g., grade, ethnicity. report the Internet user behavior to investigate the reason of addiction. examine the risk factors related to Internet addiction. Investigate the association of sleep quality and suicide attempt of Internet addicts. On the other hand, recent research in Psychology and Sociology reports a number of mental factors related to social network mental depression. Research indicates that young people with narcissistic tendencies and shyness are particularly vulnerable to addiction with OSNs. However, the above research explores various negative impacts and discusses potential reasons for Internet addiction. By contrast, this paper proposes to automatically identify SNMD patients at the early stage according to their OSN data with a novel tensor model that efficiently integrate heterogeneous data from different OSNs. Research on mental depression in online social networks receives increasing attention recently. Among them, content-based textual features are extracted from user generated information (such as

blog, social media) for sentiment analysis and topic detection. Change employs an NLP-based approach to collect and extract linguistic a content-based features from online social media to identify Borderline Personality Depression and Bipolar Depression patients. Extract the topical and linguistic features from online social media for depression patients to analyze their patterns. Analyze emotion and linguistic styles of social media data for Major Depressive Depression (MDD). However, most previous research focuses on individual behaviors and their generated textual contents but do not carefully examine the structure of social networks and potential Psychological features. Moreover, the developed schemes are not designed to handle the sparse data from multiple OSNs. In contrast, we propose a new multi-source machine learning approach, i.e., STM, to extract proxy features in Psychology for different disease that require careful examination of the OSN topologies, such as Cyber-Relationship Addiction and Net Compulsion. Our framework is built upon support vector machine, which has been widely used to analyze OSNs in many areas In addition, we present a new tensor model that not only incorporates the domain knowledge but also well estimates the missing data and avoids noise to properly handle multi source data. Caballero estimate the probability of mortality in ICU by modeling the probability of mortality as a latent state evolving over time. Propose a hierarchical learning method for event detection and forecasting by first extracting the features from different data sources and then learning via geographical multi-level model. However, the SNMD data from different OSNs may be incomplete due to the heterogeneity. For example, the profiles of users may be empty due to The SNMDD framework. privacy issue, different functions on different OSNs (e.g., game, check-in, event), etc. We propose a novel tensor-based approach to address the issues of using heterogeneous data and incorporate domain knowledge in SNMD detection.

3.2 FEASIBILITY STUDY

Preliminary investigation examines project feasibility; the likelihood the system will be useful to the organization. The main objective of the feasibility study is to test the Technical, Operational and Economical feasibility for adding new modules and debugging old running system. All system is feasible if they are unlimited resources and infinite time. There are aspects in the feasibility study portion of the preliminary investigation:

1. Technical Feasibility
2. Operation Feasibility
3. Economic Feasibility

4. SYSTEM DESIGN

4.1 PROPOSED SYSTEM

In this project, we aim to explore data mining techniques to detect three types of SNMDs [1]: 1) Cyber-Relationship (CR) Addiction, which includes the addiction to social networking, checking and messaging to the point where social relationships to virtual and online friends become more important than real-life ones with friends and families; 2) Net Compulsion (NC), which includes compulsive online social gaming or gambling, often resulting in financial and job-related problems; and 3) Information Overload (IO), which includes addictive surfing of user status and news feeds, leading to lower work productivity and fewer social interactions with families and friends offline. Accordingly, we formulate the detection of SNMD cases as a classification problem. We detect each type of SNMDs with a binary SVM. In this study, we propose a two-phase framework, called Social Network Mental Depression Detection (SNMDD). The first phase extracts various discriminative features of users, while the second phase presents a new SNMD-based tensor model to derive latent factors for training and use of classifiers built upon Transudative SVM (TSVM). Two key challenges exist in design of SNMDD: i) we are not able to directly extract mental factors like what have been done via questionnaires in Psychology and thus need new features for learning the classification models;4 ii) we aim to exploit user data logs from multiple OSNs and thus need new techniques for integrating multi-source data based on SNMD characteristics. We propose a new innovative machine learning framework called Social Network Mental Depression Identification (SNMDI) that detects potential SNMD users by designing and analyzing many important features for identifying SNMDs from OSNs, such as disinhibition, parasociality, self-disclosure, etc. Furthermore, users may behave differently on different OSNs, resulting in inaccurate SNMD detection.

4.2 SYSTEM ARCHITECTURE DIAGRAM

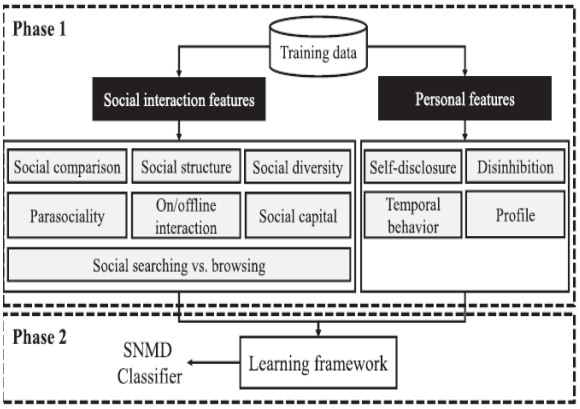


Fig 1: System Architecture Diagram

Internet Addiction has undoubtedly become the growing epidemic as the number of cases getting registered for the treatment of these mental depression due to excessive Internet Usage every year is drastically increasing. As per the latest report, this addiction has got so much to do with depression, anxiety depression, insomnia, isolation, mood swings, procrastination and many more. New terms such as Phubbing (Phone Snubbing) and Nomophobia (No Mobile Phone Phobia) have been created to describe those who cannot stop using mobile social networking apps. Conceptually, its diagnosis is a compulsive-impulsive spectrum depression that involves online and/or offline computer usage and consists of at least three subtypes: excessive gaming, sexual preoccupations, and e-mail/text messaging.

4.2 INPUT DESIGN

4.2.1. Lack of mental features

Psychological studies have shown that many mental factors are related to SNMDs, e.g., low self-esteem, loneliness. Thus, questionnaires are designed to reveal those factors for SNMD detection. Some parts of Psychology questionnaire for SNMDs are based on the subjective comparison of mental states in online and offline status, which cannot be observed from OSN logs. As it is difficult to directly observe all the factors from data collected from OSNs, psychiatrists are not able to directly assess the mental states of OSN users under the context of online SNMD detection.

4.2.2. Heavy users vs. addictive users

To detect SNMDs, an intuitive idea is to simply extract the usage (time) of a user as a feature for training SNMDI.

But, this feature is not sufficient because

1. The status of a user may be shown as “online” if she does not log out or close the social network applications on mobile phones
2. heavy users and addictive users all stay online for a long period, but heavy users do not show symptoms of anxiety or depression when they are not using social apps. To distinguish them by extracting discriminative features is critical.

4.2.3. Multi-source learning with the SNMD characteristics.

We intend to exploit user data from different OSNs in SNMDI, extracting complementary features to draw a full portrait of users while considering the SNMD characteristics into the tensor model is a challenging problem. To address the challenges, we consider a number of factors to understand the mental states of users, e.g., self-esteem and loneliness. The goal is to distinguish users with SNMDs from normal users. Two types of features are extracted to capture the social interaction behavior and personal profile of a user. It is worth noting that each individual feature cannot precisely classify all cases, as research shows that exceptions may occur. Therefore, it is necessary to exploit multiple features to effectively remove exceptions

4.3 OUTPUT DESIGN

4.3.1 EFFECTIVE FEATURES AS PROXIES TO CAPTURE THE MENTAL STATES OF USERS

A fundamental problem in text data mining is to extract meaningful structure from document streams that arrive continuously over time. Newsfeeds, messages exchanged, posts shared on an individual’s wall are all the natural examples of such streams, each characterized by topics that appear, grow in intensity for a period of time, and then fade away. A recent study has shown that social searching (actively reading news feeds from friends’ walls) creates more pleasure than social browsing (passively reading personal news feeds). This finding indicates that goal-directed activities of social searching are more likely to activate the appetitive system of a person as drug rewards do, and it is more related to SNMDs because the appetitive system is responsible for finding things in the environment that promote species survival (i.e., food, sexual mates) and thus is inclined to form addictive behaviors after several rewards. While users with

SNMDs perform social searching more frequently than non SNMDs, it is not easy to distinguish these two behaviors on social media. This example is just one such kind of a feature that could be used to analyses a user's social interaction and personal features. The new system will have many more similar features that are exploited to understand the mental status and habits of a SNMD user that considers online/offline interaction ratios, the temporal behaviors, and his self-obsessive characteristics hinting the possibility of SNMD

4.4 SYSTEM IMPLEMENTATION

Implementation is process of converting a new or revised system design into an operational one. The first task is implementation planning that is deciding on the methods and time scale to be adopted. The proper implementation involves conversion of existing clerical files to computer media and hence these files as they are get converted. Then the actual changeover from the existing system to the new system takes place. The changeover plays a vital role, which checks the developed tool for the following requirements, and then only the developed tool will be accepted by the users. The software has been checked with sample data. The changes being made are as per the user requirements and will run in parallel with the manual system to find out any discrepancies. The users also have been apprised of the ways of handling the software, as a part of training the user personnel. The implementation is the final stage and it is an important phase. It involves the individual programming system testing, user training and the operational running of developed proposed system that constitutes the application subsystems. On major task of preparing for implementation is education of users, which would really have taken place much earlier in the project when we're being involved in the investigation and design work. The implementation phase of software development is concerned with translating design specifications into source code. The user tests the developed system and changes are made according to their needs.

4.5 DATASET DESCRIPTION

Social mental depression detection train dataset input 1

Social mental depression detection test dataset input 2

In this analysis, we first apply the proposed SNMDD framework (with TSVM) on some large scale OSN datasets, i.e., FB L and IG L, to classify their users. We analyze the detected SNMD cases among the friends of an SNMD user.) The leftmost bar indicates that in FB L, among all CR users, about 45% of their friends are also CR users, which is greater than the percentage of other SNMD types. On the other hand, the 8th bar from the left in indicates that in FB L, about 59% of NC users' friends are NA (non-SNMD users). Show that, in FB L and IG L, CR and IO users have similar friend types. This is because CR and IO cases, by their nature, are similar, i.e., they are both seeking social satisfaction (e.g., relationships and information) from the OSNs. Moreover, among different SNMD cases, CR and IO users are likely to be friends with other CR and IO users. For CR users, this phenomenon has been described as "loneliness propagates" https://github.com/heroind/DATASETS

4.6 MISSING VALUE ATTRIBUTE

Unknown Value as a Regular One: An unknown (missing) value is considered as an additional attribute value. Hence, the number of values is increased by one for each attribute that depicts an unknown value in the training set.

4.7 DATA PREPROCESSSING

Data pre-processing is a data mining technique that involves transforming raw data into an understandable format. Real-world data is often incomplete, inconsistent, and/or lacking in certain behaviors or trends, and is likely to contain many errors. Data pre-processing is a proven method of resolving such issue.

Post_Weekday

Post_Hour Paid

Lifetime_Post_Total_Reach

Lifetime_Post_Total_Impressions

Lifetime_Engaged_Users

Lifetime_Post_Consumers

Lifetime_Post_Consumptions

Lifetime_Post_Impressions_by_people_who_have_liked_your_Page

Lifetime_Post_reach_by_people_who_like_your_Page

Lifetime_People_who_have_liked_your_Page_and_engaged_with_your_post comment like share

Total_Interactions class

IO User Called as Intreactive Online Users

CR-Classified rate
NC-Not Classified

5.EXPERIMENTAL RESULT

All of the variants share the following four components:

1. Excessive use, often associated with a loss of sense of time or a neglect of basic drives
2. Withdrawal, including feelings of anger, tension, and/or depression when the computer is inaccessible
3. Tolerance, including the need for better computer equipment, more software, or more hours of use
4. Negative repercussions, including arguments, lying, poor achievement, social isolation, and fatigue.

These symptoms form important diagnostic criteria for SNMDs like Cyber-Relationship Addiction, Information Overload, Net Compulsion, Cyber-Sexual and Computer Addiction. The symptoms of these depression were till now observed passively and hence the clinical intervention got delayed. Research shows that the early diagnosis of such mental depression can greatly reduce the risk.

Hence the practice of SNMD identification that relies on self-revealing of those mental factors via questionnaires in Psychology is not adopted in our proposed model as the users might try to over smart the diagnosis by answering questions dishonestly.

ALGORITHM	ACCURACY
Decision tree	87.68
SMO	86.87
Tucker	87.07
STM	90.82

Fig 2: Table Representation for Algorithm and Accuracy

5.1 PERFORMANCE EVAULATION

In the following, we first evaluate the performance of the proposed features using TSVM. We adopt Accuracy (Acc.) and Area Under Curve (AUC) for evaluation of SNMDD. Moreover, Microaveraged-F1 (Micro-F1) and Macroaveraged-F1 (Macro-F1) are also compared for multiple-label classification. Table 2 summarizes the average results and standard deviations, where the examined feature sets are denoted by self-explained labels. The results on the IG US and FB US datasets in the user study show that Duration leads to the worst performance, i.e., the results of accuracy are 34% and 36%, and the AUC are 0.362 and 0.379, respectively. Notice that the AUC function can flip the results if the calculated AUC is less than 0.5, i.e., 1-AUC.

Here, we do not filp the results to show that Duration is in fact a negative predictor in our case because Duration cannot differentiate heavy users with addictive users. Using all (All) or parts (Social or Personal) of the proposed features outperforms Duration significantly (see Table 2). All achieves the best performance (80% and 84% accuracy on the IG US and FB US datasets, respectively) because SNMDD is able to capture the various features extracted from data logs to effectively detect SNMD cases. The performance of Personal and Social are comparable, and the integrated feature set All outperforms Personal and Social by at least 15% and 16% on IG US and FB US in terms of accuracy. Since the F1 measure ignores true negatives, its magnitude is mostly determined by the number of true positives, i.e., large classes dominate small classes in micro averaging. As shown in Table 2, Micro-F1 of Duration, Social, and Personal are larger than Macro-F1 using both IG US and FB US datasets, indicating that using parts of features performs better on IO and CR (large classes) than NC. In contrast, the performance of SNMDD isalmost the same in Micro-F1 and Macro-F1, which indicates its robustness.

The results from FB US are better than those from IG US because IG US is sparser, e.g., there are no event and game posts on Instagram. After comparing the results from SNMDD with the ground truth obtained via user study, we observe that some false-positive users are detected as NC, probably because people with NC are more likely to hide their real usage time, e.g., the game logs of some people with NC are hidden. As a result, a few normal users may be incorrectly detected as NC. However,

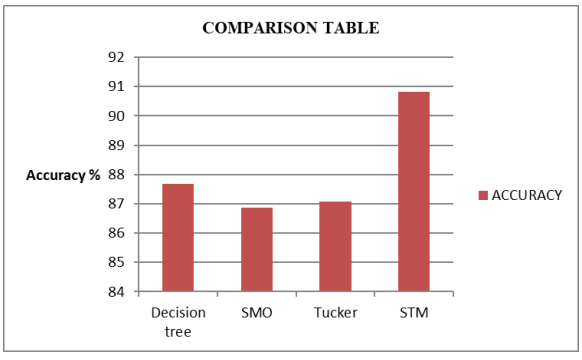


Fig 3: Graphical Representation for Performance Analysis

6. SYSTEM TESTING

System testing is the stage of implementation, which is aimed at ensuring that the system works accurately and efficiently before live operation commences. Testing is vital to the success of the system. System testing makes a logical assumption that if all the parts of the system are correct, the goal will be successfully achieved. System Testing is a type of software testing that is performed on a complete integrated system to evaluate the compliance of the system with the corresponding requirements. System testing detects defects within both the integrated units and the whole system. The result of system testing is the observed behavior of a component or a system when it is tested. System Testing is basically performed by a testing team that is independent of the development team that helps to test the quality of the system impartial. The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, subassemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

6.1 OBJECTIVES OF TESTING

Testing is the process of executing a program with the intent of finding an error.

A successful test is one that uncovers a discovered the error.

6.2 TYPES OF TESTING

A series of tests are performed for the proposed system before the system is ready for user acceptance testing.

- Unit testing
- Integration testing
- Validation testing
- Output testing
- User acceptance testing

6.3 UNIT TESTING

Unit testing focuses verification efforts on the smallest unit of software design, the module. This is also known as “module testing”. The modules are tested separately. This testing is carried out during programming stage itself. In this testing step, each module is found to be working satisfactorily as regard to the expected output from the module.

6.4 INTEGRATION TESTING

Data can be lost across an interface; one module can have an adverse effect on others; sub-functions when combined may not produce the desired major functions; integration testing is a systematic testing for constructing the program structure. While at the same time conducting to uncover errors associated within the interface? The objective is to take unit tested modules and to combine them and test it as a whole. Here correction is difficult because the vast expenses of the entire program complicate the isolation of causes. This is the integration-testing step; all the errors encountered are corrected for the next testing step.

6.5 VALIDATION TESTING

Verification testing runs the system in a simulated environment using simulated data. This simulated test is sometimes called alpha testing. This simulated test is primarily looking for errors and monitions regarding end user and decisions design specifications hat where specified in the earlier phases but not fulfilled during construction. Validation refers to the process of using software in a live environment in order to find errors. The feedback from the validation phase generally produces changes in the software to deal with errors and failures that are uncovered. Then a set of user sites is selected that puts the system in to use on a live basis. They are called beta tests. The beta test suits use the system in day-to-day activities. They process live transactions and produce normal system output. The system is live in every sense of the word; except that the users are aware they are using a system that can fail. But the transactions that are entered and persons using the system are real. Validation may continue for several months. During the course of validating the system, failure may occur, and the software will be changed. Continued use may produce additional failures and need for still more changes.

6.6 OUTPUT TESTING

After performing the validation, the next step is output testing of the proposed system, since no system could be useful if it does not produce the required output in the specified format. Asking the users about the format required by them tests the output generated or displayed by the system under consideration. Hence the output format is considered in two ways-one is on screen and another in printed format.

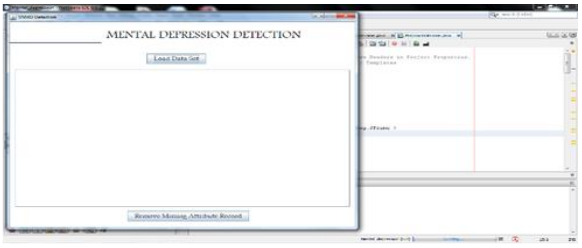
6.7 USER ACCEPTANCE TESTING

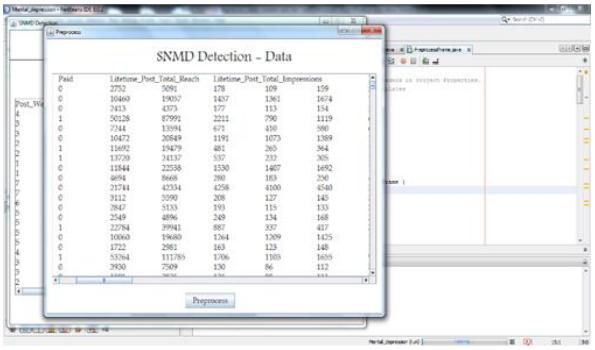
User acceptance of a system is the key factor for the success of any system. The system under consideration is tested for the user acceptance by constantly keeping in touch with the prospective system users at the time of developing and making changes whenever required. This is done in regard to the following point: An acceptance test has the objective of selling the user on the validity and reliability of the system .it verifies that the system's procedures operate to system specifications and that the integrity of important data is maintained. Performance of an acceptance test is actually the user's show. User motivation is very important for the successful performance of the system. After that a comprehensive test report is prepared. This report shows the system's tolerance, Performance range, error rate and accuracy.

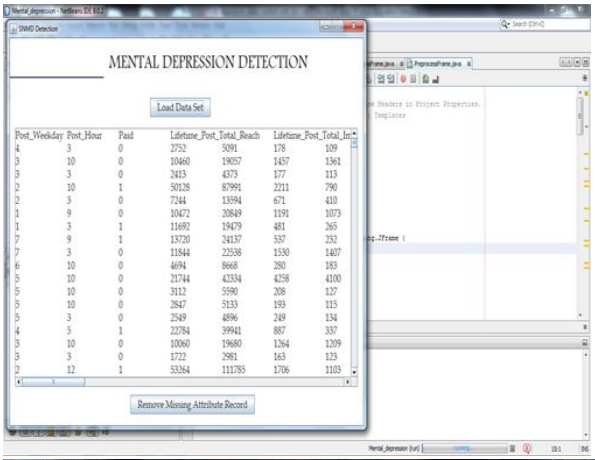
7. CONCLUSION

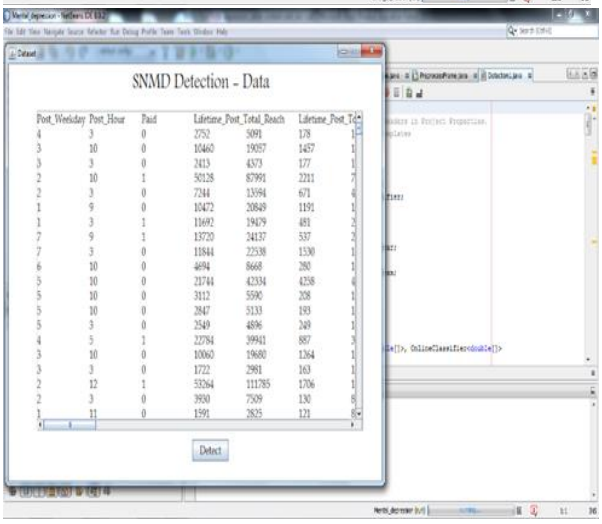
In this Project, we make an attempt to automatically identify potential online users with SNMDs. We propose an SNMDD framework that explores various features from data logs of OSNs and a new tensor technique for deriving latent features from multiple OSNs for SNMD detection. When the data from different OSNs of a user are available, the accuracy of the SNMDI is expected to improve by effectively integrating information from multiple sources for model training. A naive solution that concatenates the features from different networks may suffer from the curse of dimensionality. Accordingly, we propose an SNMD-based Tensor Model (STM) to deal with this multi-source learning problem in SNMDI. Specifically, we formulate the task as a semi-supervised classification problem to detect three types of SNMDs and the new framework can be deployed to provide an early alert for potential patients.

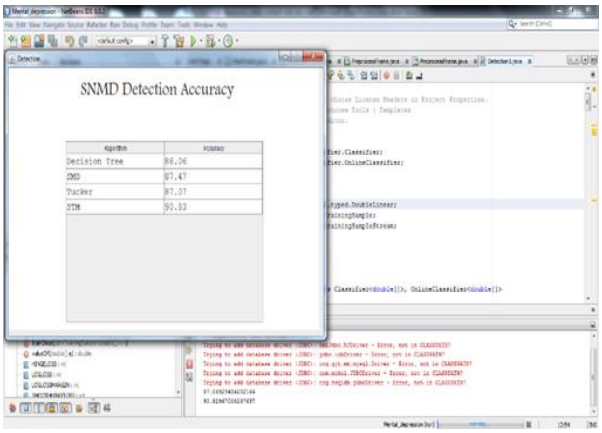
7.1 SCREENSHOTS











REFERENCES

[1] J.-W. Baek, H. Jung, and K. Chung, ‘‘Context mining based mental health model for lifecare platform,’’ Medico-Legal Update, vol. 19, no. 1, pp. 674–679, Apr. 2019.

[2] B. Mei and Y. Xu, ‘‘Multi-task least squares twin support vector machine for classification,’’ Neurocomputing, vol. 338, pp. 26–33, Apr. 2019

[3] E. Moulay, V. Léchappé, and F. Plestan, ‘‘Properties of the sign gradient descent algorithms,’’ Inf. Sci., vol. 492, pp. 29–39, Aug. 2019.

[4] G. Olmschenk, Z. Zhu, and H. Tang, ‘‘Generalizing semi-supervised generative adversarial networks to regression using feature contrasting,’’ Comput. Vis. Image Understand., vol. 186, pp. 1–12, Sep. 2019.

[5] J. C. Kim and K. Chung, ‘‘Prediction model of user physical activity using data characteristics-based long short-term memory recurrent neural networks,’’ KSII Trans. Internet Inf. Syst., vol. 13, no. 4, pp. 2060–2077, Apr. 2019.

[6] J. Li, K. Perrine, L. Wu, and C. M. Walton, ‘‘Cross-validating traffic speed measurements from probe and stationary sensors through state reconstruction,’’ Int. J. Transp. Sci. Technol., vol. 8, no. 3, pp. 290–303, Sep. 2019.

[7] J. W. Baek, K. Chung, J. Kim, and H. Jung, ‘‘Cloud-based ontology context mining using deep learning in healthcare,’’ Int. J. Innov. Technol. Exploring Eng., vol. 8, no. 8, pp. 296–300, 2019.

[8] K. Chung and J. Kim, ‘‘Activity based nutrition management model for healthcare using similar group analysis,’’ Technol. Health Care, vol. 27, no. 5, pp. 473–485, Sep. 2019.

[9] M. J. Shafiee, B. Chwyl, F. Li, R. Chen, M. Karg, C. Scharfenberger, and A. Wong, ‘‘StressedNets: Efficient feature representations via stressinduced evolutionary synthesis of deep neural networks,’’ Neurocomputing, vol. 352, pp. 93–105, Aug. 2019.

[10] Y. Chen, J. Jiang, and A. Narayan, ‘‘A robust error estimator and a residualfree error indicator for reduced basis methods,’’ Comput. Math. Appl., vol. 77, no. 7, pp. 1963–1979, Apr. 2019.