



CLASSIFICATION OF DIABETES MELLITUS IN HUMANS USING DEEP ARTIFICIAL NEURAL NETWORK

Dr.T.K. Thivakaran,

Professor and Program Head, Department of CSE, Presidency University, Itgalpura, Bangalore,
India Email: tktpresidency4@gmail.com

Dr.D.Sasikumar,

Professor, Dept of CSE, Sphoorthy Engineering College, Hyderabad, Telangana, India Email:
drsasikumar@sphoorthyengg.ac.in

Dr.M.Nagarajan,

Associate Professor, Department of ECE, Bharath Institute of Higher Education and Research,
Chennai, India Email: nagarajmohan@yahoo.com

Dr.A.SAIRAM,

Associate Professor, Department of Computer Science and Engineering,
Thirumalai Engineering College, India Email: Dr.a.sairaam@gmail.com

Corresponding Author: Dr. T. K. Thivakaran, Email: tktpresidency4@gmail.com

10070

Abstract

This paper aims to make use of the information that has been discussed in a community for people who have type 2 diabetes in order to assist in directing the responsible utilization of the medicine sitagliptin. The research make use of word frequency data and an exploratory study phase that makes use of deep neural network-based self-organizing maps (SOMs). Lastly, we carry out network analysis by making use of customized graph theory approaches in order to locate user groups (modules) as well as potential users. Simulation is conducted to test the efficacy of the model against various performance metrics. The results of simulation shows that the proposed method achieves higher accuracy than other methods.

Keywords: Type 2 Diabetes, Classification, Neural Networks

DOI Number: 10.14704/NQ.2022.20.6.NQ22981

NeuroQuantology2022;20(6): 10070-10080

1. Introduction

Users are able to communicate with one another on these platforms, increasing the likelihood that they may become friends, admirers, and supporters of the content produced by one another. Conversations that take place online have developed into an increasingly significant part of the human experience as a direct result of the proliferation of social media [1].

When more individuals utilize social media, there are more opportunities to analyze the behaviors of big groups of people, learn about individual users, and better advertising. Take,

for instance, the practice of putting adverts on various platforms that are associated with social media. The advertising that takes place on social media platforms has given rise to a number of problematic situations. The formation of a model that classifies user profiles in accordance with characteristics such as geographic location, educational attainment, and gender is a strategy that is utilized frequently [2].

This is one perspective to take on the situation, as it can be seen from the point of view of a classification problem. The primary objective of this initiative is to create relevant



advertising groups by categorizing clients according to their preferences. In order to find a solution to the classification issue, there is only a constrained amount of data that can be categorized and compiled. It is possible to infer a person likely interests based on their online behavior, such as when they click on an advertising, make a purchase, or mention a hobby or interest on their profile [3]. This can be done by looking at patterns of clicks, purchases, and mentions of hobbies and interests. On the other hand, the label itself does not provide a significant amount of information about the product. We investigate the various ways in which the vast amounts of data that are currently available can be used to categorize activities that take place on social media platforms. This is done while taking into account the social networks of individuals as well as the information that these individuals have about their relationships [4].

The study of social structures through the prism of network topology was the subject of research carried out in the academic discipline of social sciences. Questionnaires are a tried-and-true method of data collection, particularly useful in the social sciences. When filling out a questionnaire, participants are frequently requested to provide details on their experiences communicating with a diverse range of individuals in a variety of environments. After the responses have been compiled, a network can be constructed with the nodes of the network standing in for the many individuals who have been in contact with one another and playing the role of connecting points between the various other nodes [5]. A social network analysis conducted using this form of data collecting can at most include a few hundred people, which is much too small to make any conclusions that have any meaningful importance whatsoever. Many distinct kinds of connections can be found operating inside the confines of a single network. The information provided by these linkages enables research to be carried out on topics such as social norms and roles, as well as the dynamics of groups [6].

Relationships of many different types play a role in the formation of links within the enormous networks that are made possible by social media. A user has the opportunity to meet others online with whom they share familial and social ties, educational or professional interests, a common pastime or passion, or any combination of these things. The social network provides an inherently diverse range of linkages, and these linkages serve as a representation of the many various forms of user relationships. However, when a social network is built, the underlying reasons for the connections between individuals and the nature of their relationships are sometimes obscured [7].

Rarely is this variability taken into account by the many solutions to classification problems that are now accessible, particularly when such solutions involve connected data. This suggests that current methods address inhomogeneous connections in a way that is typically consistent across a wide range of contexts and applications. Using a network that includes people from a wide variety of backgrounds will almost certainly lead to the occurrence of unintended consequences [8].

The main contribution of the work involves the following: The manner in which user social media connections influence the user beliefs and behaviours are taken into account by our research. In contrast to studies that have been published in the past, our initial step was to identify notable users. Second, we have made enhancements to our system that automatically assigns positive and negative implications to words in posts based on the context in which the words are revealed. Thirdly, we make use of word frequency data and an exploratory study phase that makes use of deep neural network-based self-organizing maps (SOMs). Lastly, we carry out network analysis by making use of customized graph theory approaches in order to locate user groups (modules) as well as potential users.

2. Related works

The process of classifying text begins with the compilation of a set of input documents and an established grouping of categories that have previously been determined. After then,

10071

it makes an effort, using the documents that were provided as input, to fit each document into one or more of the classes. Due to the fact that this research is widely recognized as a significant milestone achievement in the domains of text mining and text learning, a significant amount of attention was given to it. The categorization can be split down even further into binary, multi-class, and multi-label categories depending on how the findings of the labeling are formulated [9].

The classifying texts into binary categories or many classes is an important one in the field of machine learning. In the past, this challenge has been tackled in great depth for traditionally large volumes. Because of the recent increase in the amount of text that has been exchanged on social media platforms, there has been a considerable amount of focus placed on the categorization of short texts, in particular the classification of social text streams. The application of multiple hierarchically ordered labels (HMC) in document classification presents yet another obstacle for those involved in text classification [10].

Over the course of the past several years, there has been a substantial amount of interest among the academic research community about the categorization of brief writings. The modern literary canon includes shorter works that address the issue of information scarcity by making use of knowledge obtained from other sources. This is done in order to address the fact that there is a shortage of information. Both web-based methodology and topic-based approaches can be utilized in the process of defining these strategies as viable options [11].

Web-based search algorithms interpret every short text as a search engine query and make use of the external knowledge that may be learned from the results of searches in order to increase the efficiency with which they classify short texts. the performance of web-based search engines can be improved in terms of categorizing brief amounts of text. Because of inefficiencies and limitations in their ability to scale, techniques of this kind are not suitable for use in the data-rich environment in which we currently operate.

This is because of the nature of the data. It has been suggested in a number of different research that putting together a very large corpus of text is the best way to improve the effectiveness of categorization systems [12].

The performance of web-based search engines can be improved in terms of categorizing brief amounts of text. Because of inefficiencies and limitations in their ability to scale, techniques of this kind are not suitable for use in the data-rich environment in which we currently operate. This is because of the nature of the data. It has been suggested in a number of different research that putting together a very large corpus of text is the best way to improve the effectiveness of categorization systems [13].

In terms of methodological techniques that are focused on topics, the authors [14] extract the theme distributions from the Wikipedia dump using the LDA model. A approach that is both successful and efficient for the extraction of numerous granularities of latent issues from large-scale external training sets.

In addition to these two strategies, other other approaches were applied during the process. The results of categorization were found to be much enhanced when basic text was converted into entities. The authors in [15] are able to develop a short text classification that makes use of the fact that some test document instances may have common discriminating texts with the training set because they linked the info track. In doing so, they are able to make use of the fact that are able to develop a short text classification. They were able to use this to their advantage by taking advantage of the fact that some instances of the test document shared common discriminating terms. Extremely few of the works that are currently available on classifying brief texts take a streaming environment into consideration, and none of them approach the problem from a multi-label hierarchical point of view. Streaming environments are becoming more and more prevalent in today online world.

3. Proposed Method

One of the social networks that is dedicated to type 2 diabetes mellitus has been found to have one of the biggest numbers of users, as

10072

indicated by the preliminary findings. We contrasted the findings that we received after searching through the most prominent diabetes-related discussion groups on Facebook, Twitter, and WebMD.

A database of patient pharmaceuticals and medical equipment was built and searched for in each forum in order to determine which treatments are of the most interest to users. This was done so that the most popular remedies could be identified. Sitagliptin was found to be the medicine that is spoken about the most frequently on message boards and in online communities after a compilation of online communities and message boards was conducted and a search for sitagliptin was conducted in each forum. It has been discovered that the medicine known as sitagliptin is the one that is discussed the most frequently on discussion boards and in online forums.

When the search query on the DiabetesDaily.com message board was entered with the name of the generic medicine Sitagliptin, the results were not as

comprehensive as they were when the name of the brand drug Januvia was used instead (92 outcomes). When searching for Januvia in 2007, 713 results were located on various websites.

Initial Text Mining And Preprocessing

After that, we determined the texts that were used the most frequently in the posts by utilizing a decision tree that had been updated in the software Rapidminer (www.rapidminer.com). In order to make room for the characters that are more desirable, this tree has gotten rid of the characters that are more common but less desirable. This was done in order to keep the tree organized.

The completed project consisted of compiling a wordlist of social media posts and including TF-IDF values underneath each entry. Words are assigned the labels positive or negative, and in order to further categorize them, the frequency-inverse document frequency (TFIDF) scheme is inserted into each vector, also known as a posting.

10073

$$w_{t,d} = \begin{cases} \log(tf_{t,d} + 1) \log \frac{n}{x_t} & \text{if } tf_{t,d} \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

where

$tf_{t,d}$ is considered as the word frequency t in a document d ,

n is considered as the total number of documents and

x_t is considered as the total number of documents with the occurrence of word t .

Text Tagging

Following an analysis of the social media posts, the highest TF-IDF scores were found to belong to the following words: After that, these words were put into one of two categories: either positive or negative, depending on how strongly they were connected to the concepts of context positivity and context negativity, respectively. A reduction in the number of words that were quite comparable to one another that had been identified as a result. This was achieved

both manually (by comparing the words with Merriam-Webster) and automatically (by using synonym database tools such as the Thesaurus Synonym Database and Google synonymous search finder by typing the Greek letter chi) to achieve the desired results. Both of these approaches were applied.

The word list in Table 1 was reduced using the following techniques, and the chart that was produced as a result of this procedure can be found lower down the page. The words that are most strongly connected to either positive or negative emotions have been italicized in this list. In order to arrive at this decision, not only the frequency of the terms but also their context within the network and the messages themselves were taken into consideration.

Table 1: The Wordlist of Final Post-Analysis

Positive	Negative
Good	Worry
Doesn't_p	Friendly_n
Wasted_p	Like_n
Wrong_p	Great_n
Fine	Not
Right	Problems
Didn't_p	Hope_n
Isn't_p	Happily_n
Easier	Havent
Yes	Pain
Hopefully	Never
Wouldn't_p	Improvement_n

10074

In addition, we got rid of any and all words that were used in the primary text fewer than ten times altogether. We were able to construct a standardized collection of measures as a result of this by removing the data points that were statistically unimportant as part of the process. The complete list of 28 words is presented here in Table 3.1 for your perusal. It is made up of 14 words that have not been tagged, in addition to 14 terms that have not been tagged but have been labeled as having a bad connotation (13 untagged words).

Before the data were gathered for exploratory analysis utilizing self-organization maps, a manual rating of positive, negative, or neutral was assigned to each item depending on the user perspective and assigned before the data were collected. This evidence can be used to determine whether or not the accuracy of the data collected by hand labeling has been maintained.

Classification Using SOM

An example of an artificial neural network that can be used to cluster high-dimensional data in a way that yields localized representations is known as a self-organizing map, which is also sometimes referred to as a SOM (Kohonen 2000). A self-organizing map, also known as a SOM, is a type of network in which each neuron represents a cluster of weight vectors. The space that is formed as a result of this (the projected input data) is represented by a neural layer in the network. The values of the weight vectors that are used to construct a cluster provide a direct reflection of the contents of that cluster. The data that was gathered by the network and presented by the SOM can be used to match up related data vectors with the appropriate neurons. This can be done by using the data that was acquired by the network.

Because of the significant contribution made by the high-dimensional data analysis, we were able to successfully implement the self-ordering map (SOM) due to the aesthetic benefits and high-level capacity it offers. This

was made possible by the significant contribution made by the high-dimensional data analysis. In the process of identifying clusters, Bonato et al. (2003) demonstrated how vector quantization techniques can reduce the amount of space consumed by the features without ignoring any of the essential information. This was accomplished by reducing the amount of space occupied by the features.

During the training process, the network receives new data to be ingested into it. After that, the method will look through the network for the node that possesses the weight vector that is the most similar to its own, and it will give that node the data vector. Following this, the values of the neurons within that node and those surrounding it will be altered so that they are reflective of the newly inserted value. Even though it is to a lesser extent, the weight vectors of newly formed neurons travel farther away from the weight vectors of neurons that have undergone changes. This method is repeated with each and every input vector as many times as is required up until the point where the convergence requirements are satisfied. As the project draws to a close, a map that has only two dimensions will be created.

We have investigated, with the help of Matlab SOM toolbox, whether or not particular vectors cluster together in accordance with the categories found in the revised word list that we have. For the purposes of internal validation, when we were training our SOM on maps of varying sizes, we purposely included quantization errors as well as topographical imperfections. Calculating a sample error involves first determining the average distance between each input vector and the neuron that corresponds to it the best, and then dividing this value by the total number of input vectors.

Word list vectors were moved to the SOM and emergent clusters so that more investigation could be conducted into the connections that exist between positive and negative characteristics. We took out of circulation any groups that had three posts or fewer and had no relevant search terms in their profiles. A

tally was performed to establish the frequency of occurrence of each individual term within the clusters that were retained.

The subgroups of the map were visually identified, and we evaluated the correlations between the images and the texts in order to determine whether or not they were representative of consumer pleasure (or dissatisfaction) with Sitagliptin.

SOM

Figure 1 presents a high-level overview of the training process, in addition to several examples that illustrate various aspects of it. Imagine that the various sections of the brain that are responsible for processing data are separate parts of the brain that need to learn specific abilities in order to be able to do certain tasks effectively.

The research focuses on two major features: the first is the learning rate of the SOM network, and the second is the gaining neighborhood. Both of these aspects are gaining neighborhood and learning rate. According to the statistics, the SOM network is able to learn more effectively when both its acquiring neighbor domain, $h_{c,j}(t)$, and its learning rate, $a(t)$, are slowed down. This results in the network taking a greater amount of time to learn new information, but it does so in a more effective manner. In addition to this, the SOM network usually leads to a higher level of self-organization as a result of its use.

However, in order to attain the superior $a(t)$ and the winning $h_{c,j}(t)$ next-door domain, one must first adjust the learning rate of the SOM network. How exactly does this process take place? However, the only way to arrive at superior experimental settings is by experience or experimentation; there is no valid theoretical guidance that can be used. There is a severe flaw in the SOM neural network that has to be corrected as soon as possible.

After the weight vectors have been fine-tuned to their optimal values, the SOM neural network will be outfitted with reward neurons when that time comes. This will take place after the weight vectors have been fine-tuned. The process of optimizing the adjustment can be done on a local level with a

10075

good amount of ease since neurons that have substantial error rates in their performance can be ignored. This makes the procedure more efficient overall. There is not only no theoretical basis for determining the size of

the domain that will be awarded to the neighbor who is the winner, but there is also no evidence to imply that larger domains are preferred in any circumstance.

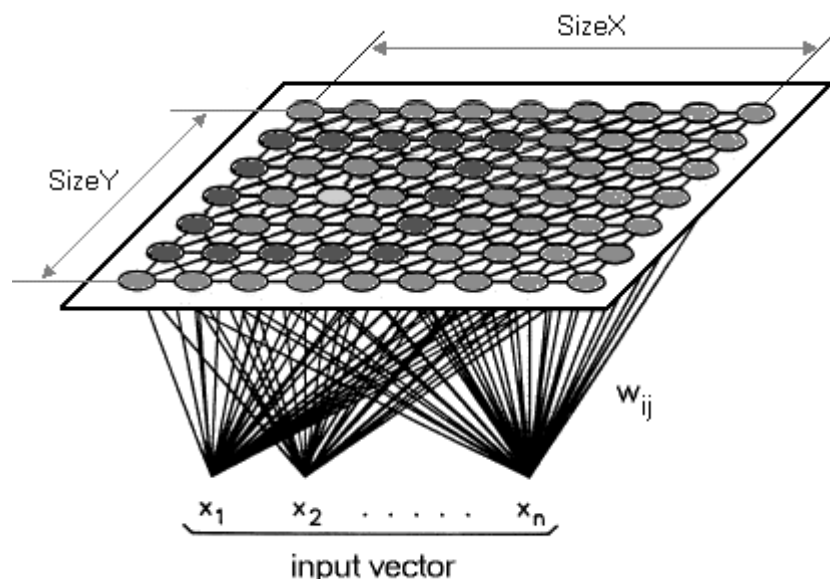


Figure 1: SOM neural network

If the units are connected in such a way that some of the work can be learned from the (predefined) neighbors in the unit, then the system will also grow slower as a result of this configuration. The capacities of neighboring units are equivalent, and the system as a whole is able to adapt to new demands in a way that is both gradual and seamless. This is all done in accordance with the self-organizing map (SOM) principle.

When unsupervised learning is carried out, the representations are arranged in such a way as to lay a focus on the characteristics that they have in common with one another. This skill contributes to the SOM power to organize huge data sets, such as document sets, which can be relatively difficult to perform manually and contributes to the SOM capacity to organize large data sets. The map is tasked with expressing statistical data that is provided by vectors, and it is produced in the form of a regular grid of units for the most fundamental type of SMEs.

When it comes to the actual representation of the data, each map unit has its own unique model for the storage of vectors. The map is

rendered worthless as a consequence of the learning process, which, over the course of time, results in the model vectors moving. As a direct result of this, the model vectors in the data space are regressed in a manner that is non-linear in nature.

In order for the training procedure for the SOM to work, the weight vectors must be inverted before they can be made to pass through the data cloud. This assists in organizing the map and producing equal weight vectors for neurons that are located in close proximity to one another on the grid. Samples are displayed on the map one at a time in the conventional sequential training technique, after which the algorithm makes progressive adjustments to the weight vectors in their respective directions. During batch training, the entire SOM is shown with the dataset, and the newly trained weight vectors are determined by computing the weighted averages of the data vectors. During batch training, the dataset is also supplied with the complete SOM as in Figure 3.

Let pretend for a moment that the system is taking in information in the form of vectors

$X(t)$ that have a dimension of d , and that k neurons are going to be in charge of producing output. The neural network initialization has a significant influence on many aspects of the network, including the

$$\|X(t) - w_c\| = \min_j \|X(t) - w_j\|, j = 1, 2, \dots, k$$

Give the neural network a training sample, which will be indicated by the notation $X(t)$, and then choose a winning neuron, which will be denoted by the notation c , by using a competition strategy that is based on the minimum feasible Euclidean distance.

$$w_j(n+1) = w_j(n) + a(t)h_{c,j}(t)(X - w_j(n))$$

where, t - c -domain function of the local neuron, which is represented by $h_{c,j}(t)$, as well as the learning rates, which are denoted by a , are displayed for our perusal and evaluation $a(t)$. This process is repeated as many times as required until there is no longer any variation in the value or condition of the connection weight vectors. This can take anywhere from a few minutes to several hours.

4. Results and Discussions

In this section, a comparison is made between the Deep Neural Network SOM that was proposed and the classification methods that

learning rate, the size of the network, the number of studies, and the value of the initialized link weight vectors. These factors are all affected by the value of the initialized link weight vectors.

Find the neighbor domain that has the highest connection weight, $h_{c,j}(t)$ and after that, modify all of the scalars that represent the connection weight between neighbor domains.

already exist. The performance of each method is evaluated using a range of metrics. Accuracy, G-mean, recall, F-measure, and percentage error are only few of the metrics that can be used. Other metrics include the following: The following is a list of some of the ways in which the old approaches to SOM differ from the Deep Neural Network SOM method that has been developed: Network Programmable Self-Organizing Map (NPSOM), Neural Network based Self-Organizing Map (NN-SOM) and Bayesian Networks based SOM (BN-SOM).

10077

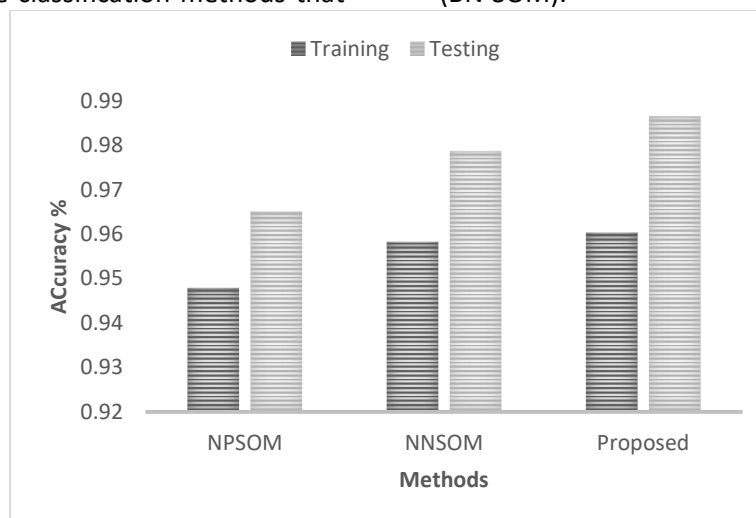


Figure 2: Accuracy

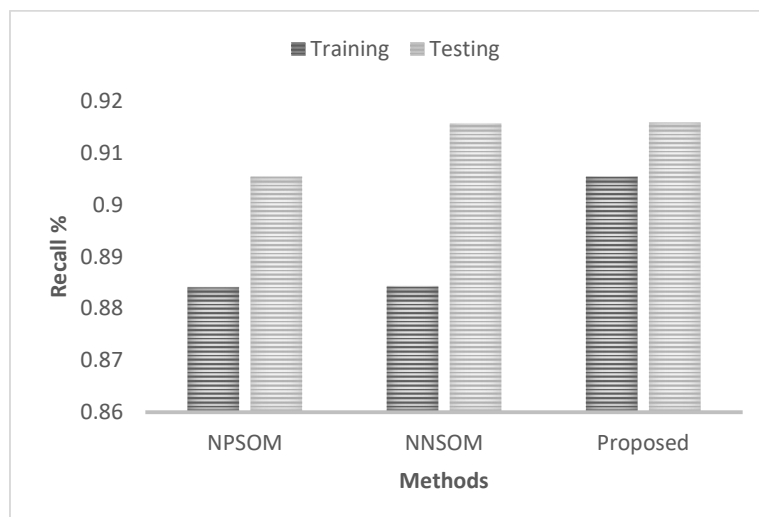


Figure 3: Recall

10078

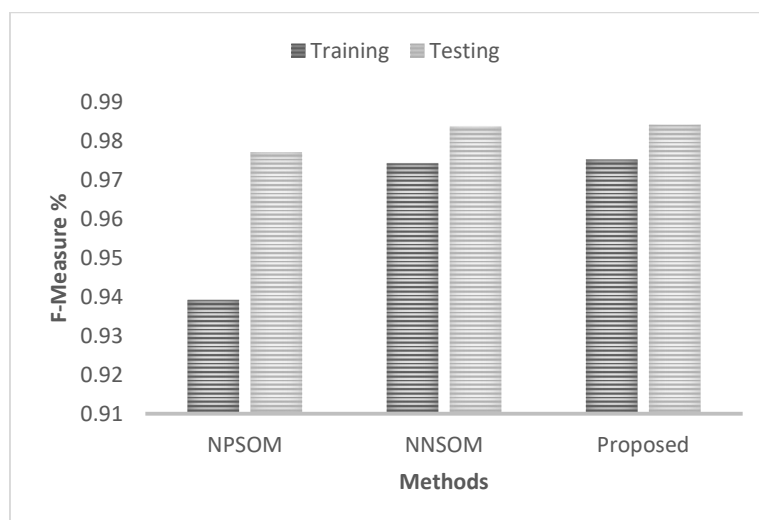


Figure 4: F-Measure

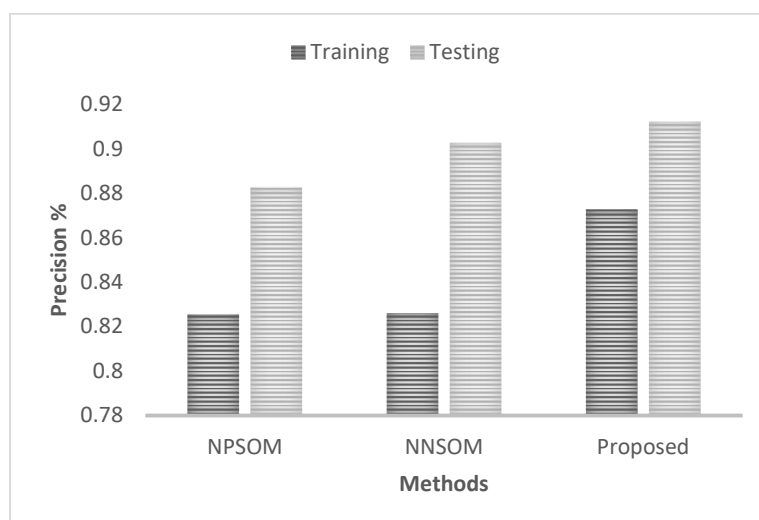
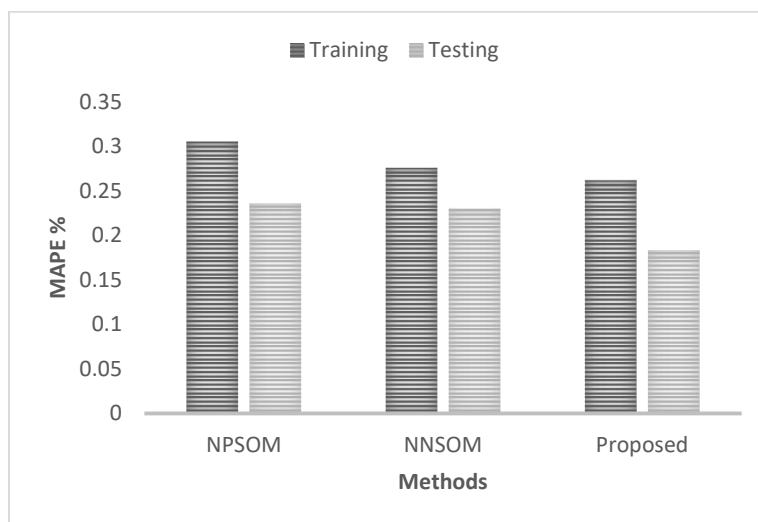


Figure 5: Precision



10079

Figure 6: MAPE

From the results of simulation shown in Figure 2 to Figure 6, shows that the proposed method achieves higher rate of classification accuracy, precision, recall and f-measure in classifying the texts than the existing methods.

5. Conclusions

In this paper, we apply the used frequently used words for SOMs to employ the exploratory purposes. Finally, in order to distinguish between preexisting user groups (modules) and prospective new users, we conduct network analysis utilizing specialized graph theory methods. To give suggestions for the effective and secure utilization of sitagliptin, we simulate the model's performance in relation to a variety of key markers of type 2 diabetes in patients. Data on how often certain words are used is used, and SOMs are also employed for exploratory purposes. The effectiveness of the model is measured against a variety of metrics using simulation. The simulation results demonstrate that the suggested method outperforms the state-of-the-art methods. The analysis of our data revealed that there are positive and negative perspectives on the efficacy and side effects of sitagliptin. Extensive follow-up studies confirmed these hypotheses as sound scientific grounding.

References

- [1] Sungheetha, A., & Sharma, R. (2021). Design an early detection and classification for diabetic retinopathy by deep feature extraction based convolution neural network. *Journal of Trends in Computer Science and Smart technology (TCSST)*, 3(02), 81-94.
- [2] Kamalraj, R., Neelakandan, S., Kumar, M. R., Rao, V. C. S., Anand, R., & Singh, H. (2021). Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm. *Measurement*, 183, 109804.
- [3] Jangir, S. K., Joshi, N., Kumar, M., Choubey, D. K., Singh, S., & Verma, M. (2021). Functional link convolutional neural network for the classification of diabetes mellitus. *International Journal for Numerical Methods in Biomedical Engineering*, 37(8), e3496.
- [4] Jader, R., & Aminifar, S. (2022). Fast and accurate artificial neural network model for diabetes recognition. *NeuroQuantology*, 20(10), 2187-2196.
- [5] Sarki, R., Ahmed, K., Wang, H., Zhang, Y., & Wang, K. (2022). Convolutional neural network for multi-class classification of diabetic eye

- disease. *EAI Endorsed Transactions on Scalable Information Systems*, e15-e15.
- [6] Thaiyalnayaki, K. (2021). Classification of diabetes using deep learning and svm techniques. *International Journal of Current Research and Review*, 13(01), 146.
- [7] Nagi, A. T., Awan, M. J., Javed, R., & Ayesha, N. (2021, April). A Comparison of Two-Stage Classifier Algorithm with Ensemble Techniques On Detection of Diabetic Retinopathy. In *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)* (pp. 212-215). IEEE.
- [8] García-Ordás, M. T., Benavides, C., Benítez-Andrades, J. A., Alaiz-Moretón, H., & García-Rodríguez, I. (2021). Diabetes detection using deep learning techniques with oversampling and feature augmentation. *Computer Methods and Programs in Biomedicine*, 202, 105968.
- [9] Chen, H. C., Liau, B. Y., Lin, C. Y., Akbari, V. B. H., Lung, C. W., & Jan, Y. K. (2021). Estimation of various walking intensities based on wearable plantar pressure sensors using artificial neural networks. *Sensors*, 21(19), 6513.
- [10] Das, S. K., Roy, P., & Mishra, A. K. (2021). Deep learning techniques dealing with diabetes mellitus: a comprehensive study. In *Health informatics: A computational perspective in healthcare* (pp. 295-323). Springer, Singapore.
- [11] Goel, S., Gupta, S., Panwar, A., Kumar, S., Verma, M., Bourouis, S., & Ullah, M. A. (2021). Deep Learning Approach for Stages of Severity Classification in Diabetic Retinopathy Using Color Fundus Retinal Images. *Mathematical Problems in Engineering*, 2021.
- [12] Valizadeh, A., Jafarzadeh Ghouschi, S., Ranjbarzadeh, R., & Pourasad, Y. (2021). Presentation of a segmentation method for a diabetic retinopathy patient's fundus region detection using a convolutional neural network. *Computational Intelligence and Neuroscience*, 2021.
- [13] Anisuzzaman, D. M., Patel, Y., Rostami, B., Niezgoda, J., Gopalakrishnan, S., & Yu, Z. (2021). Multi-modal Wound Classification using Wound Image and Location by Deep Neural Network. *arXiv preprint arXiv:2109.06969*.
- [14] Bukhari, M. M., Alkhamees, B. F., Hussain, S., Gumaei, A., Assiri, A., & Ullah, S. S. (2021). An improved artificial neural network model for effective diabetes prediction. *Complexity*, 2021.
- [15] Kumar, R., & Gupta, M. (2022). Optical coherence tomography image based eye disease detection using deep convolutional neural network. *Health Information Science and Systems*, 10(1), 1-16.

10080