

Artificial Intelligence Based Study on Analyzing of Habits and with History of Diseases of Patients for Prediction of Recurrence of Disease Due to COVID-19

Shawni Dutta¹ and Prof. Samir Kumar Bandyopadhyay²

¹Assistant Professor, Department of Computer Science, The Bhawanipur Education Society College, Kolkata, INDIA

²Professor, Academic Advisor, Department of Computer Science, The Bhawanipur Education Society College, Kolkata, INDIA

¹Corresponding Author: 1954samir@gmail.com

ABSTRACT

A patient will visit physicians when he/she feels ill. This illness is not for COVID-19 but it is a general tendency of human being to visit doctor probably it cannot be controlled by general drug. When a patient comes to a doctor, the doctor examines him/her after knowing his/her problem. The physician always asks him/her about some questions related to him/her daily life. For example, if a young male patient comes to a doctor with a symptom of fever and cough, the first question doctor asked him that he has a habit of smoking. Then doctor asks him whether this type of symptom appeared often to him previously or not. If the answers of both questions are yes, then the first one is habit and the second one is that he may suffering from some serious disease or a disease due to the weather. The aim of this paper is to consider habit of the patient as well as he/she has been affected by a critical disease. This information is used to build a model that will predict whether there is any possibility of his/her being affected by COVID-19.

This research work contributes to tackle the pandemic situation occurred due to Corona Virus Infectious Disease, 2019 (Covid-19). Outbreak of this disease happens based on numerous factors such as past health records and habits of patients. Health records include diabetes tendency, cardiovascular disease existence, pregnancy, asthma, hypertension, pneumonia; chronic renal disease may contribute to this disease occurrence. Past lifestyles such as tobacco, alcohol consumption may be analyzed.

A deep learning based framework is investigated to verify the relationship between past health records, habits of patients and covid-19 occurrence. A stacked Gated Recurrent Unit (GRU) based model is proposed in this paper that identifies whether a patient can be infected by this disease or not. The proposed predictive system is compared against existing benchmark Machine Learning classifiers such as Support Vector Machine (SVM) and Decision Tree (DT).

Keywords-- Covid-19 Test Results, Prediction, Habits, Health Records, Deep Learning

I. INTRODUCTION

The World Health Organization (WHO) named Corona Virus Infectious Disease as Covid-19. Outbreak of this disease occurred in Wuhan, Hubei Province, China, in December 2019. WHO on March 11, 2020

declared COVID-19 as 'Pandemic' [1]. During the history of mankind there had been various pandemics in the globe. Some are still disastrous than the others to the humans. Recently, humans in world are facing an unknown virus- COVID-19. It has different in nature than other virus and scientists are still searching for vaccine for this new virus. It is spreading fast in the world and all countries are facing this life threat disease. Almost normal functioning of human lives is in danger and authorities impose restrictions often in the name of lockdown in their territories.

The negative impact of tobacco use on lung health and its causal association with a plethora of respiratory diseases have been highlighted [2]. The use of tobacco is detrimental to the immune system and it is more vulnerable to infectious diseases. Alcohol consumption is likely to increase the health risks if a person made it a habit. It will not destroy the virus. It will infect mouth and throat. A couple of peoples think that and it will give protection against COVID-19. Practically it is true that a section of persons affected by COVID-19 who are drinking alcohol due to the stress in their profession. Alcohol has a bad effect on immune system and will not stimulate immunity and virus resistance [2-3]. COVID-19 symptoms are heterogeneous in nature ranging from mild flu-like symptoms to acute respiratory distress syndrome, multiple organ failure and death. The significant predictors of morbidity and mortality are age, diabetes patients and other comorbidities. The United States CDC has estimated that the incubation period of symptoms will usually develop within 2–14 days after exposure of COVID-19 affected peoples. So fourteen days has been the time defined as quarantine period [1, 4].

For COVID-19 it is required to take preventive measures by all and it is the first line of defence. The guidelines of WHO are to be followed for protecting yourself from COVID-19 [5]. These guidelines are:

1. Maintain at least 3 feet distance between yourself and anyone who are meeting or talking with you.
2. It is required to avoid close contact with sick people.
3. If possible then try to work from home.
4. It is necessary to clean and disinfect frequently touched objects and surfaces.
5. Soap and water is used to wash your hands often.

6. If you are sick, then stay home.
7. If you are going outside of your home, then use face mask and disposable gloves.
8. Take care of children without assuming any risk to children from COVID-19.
9. Avoided crowded place since it may affect you from coronavirus.
10. Touching your eyes, nose, and mouth as far as possible.
11. Visit to relatives and friends are less as far as possible
12. Before re-entering your home, take your shoes off and leave them outside, or spray them with a disinfectant. Remove and discard the mask and disposable gloves.
13. Disinfect your hands at regular interval when you are in home as well as in office.
14. Dispose of the tissue safely after cover your cough or sneeze with a tissue.
15. Clean and disinfect frequently-touched

COVID-19 creates these new habits to humans and humans have to follow these until and unless proper tested vaccine for COVID-19 will be discovered.

The medical conditions of human are at increased risk for severe illness from COVID-19 and it can be happened for any age. COVID-19 may introduce some changes of some diseases as per study made by the Physicians/Scientists. These are- Chronic kidney disease, Cancer, chronic obstructive pulmonary disease (since respiration problem is seen in all COVID-19 patients), Serious heart conditions, Asthma, Diabetes, High blood pressure, Liver disease and many more [6].

As discussed above, it can be inferred that smoking, alcohol consumption, diabetes, hypertension tendency can lead to COVID-19 occurrence in patients. Thus, it is necessary to consider these parameters along with other considerable factors while predicting patients with positive or negative symptoms. This prediction will assist medical experts to combat with this disease in an efficient manner. This paper focuses on predicting the outcome of covid-19 results either as positive or negative. For acquiring the prediction results, data mining techniques are applied on past habits and medical history of patients. Data mining refers to set of prominent techniques which are applied on larger and/or complex database for inferring undiscovered patterns [7]. This step will provide a knowledge base for accompanying informed decision-making process. Deep Learning (DL), a subfield of Machine learning (ML) techniques are explored in this context so that future predictions can be obtained while uncovering relevant data patterns. Analysing past habits as well as health records of patients, it is possible to simulate an automated tool that predicts positive or negative covid-19 results.

Recurrent Neural Network (RNN) [8] follows deep learning technique which is designed as proposed method by this paper. More specifically, this paper implements multiple Gated Recurrent Unit (GRU) [9] layers and stacks into a single platform as an automated tool. It is to be noted that, GRU is an improved version RNN. Implementation of this deep model necessitates

adjustment of hyper-parameters so that maximized performance can be obtained. This deep model is compared with traditional ML classifiers such as Decision Tree (DT) [10], Support Vector Machine (SVM) [11].

The contribution made by this study for covid-19 disease classification includes:

- Processing of healthcare and past habit data using deep learning techniques instead of the traditional healthcare system to identify COVID infected person.
- This work implements stacked-Bi-GRU model to be applied for processing patient data and identifies likelihood Covid-19 infection. This model is adjusted by fine-tuning its hyper-parameters for maximising the prediction performance. Other ML algorithms such as SVM and DT are used to compare the proposed implementation for detecting covid-19 infection possibility.

II. RELATED WORKS

Global economy and day to day life are affected by COVID-19. It is slowing down the global economy. Peoples in globe are either sick or are being killed due to the spread of this disease. Common symptoms of this viral infection are fever, cold, cough, bone pain and breathing problems, and ultimately leading to pneumonia. The new viral disease is affecting humans for the first time. The vaccines are not yet discovered. The emphasis is on taking extensive precautions such as extensive hygiene protocol, social distancing, and wearing of masks, and so on. It is now spreading exponentially region wise. Banning gatherings of people to the spread and break the exponential curve are trying by countries [12-13]. Lockdown by many countries and enforcing strict quarantine are the control parameters to avoid the spread of the havoc of this highly communicable disease. It is now essential to identify the disease at an early stage for controlling the spread of the virus because it is very rapidly spreads from person to person. Manufacturing of the products are slowed down by many countries [14-15]. Various industries and sectors are affected by the cause of this disease. COVID-19 creates significant knock-on effects on the daily life of citizens as well as about the global economy.

Researches on the psychological behaviours and habits of humans have doing at random due to the current pandemic [16]. Recent researches have directed towards the need of mental healthcare by health professionals [17]. It is also emphasized how implementing psychological support systems during the pandemic for ensuring emotional stability. It is now needed to understand the nature of the psychological consequences stemming from this newly emerging disease and initial research efforts are primarily linked to mental health consequences that individuals may experience as a result

of isolation and quarantine. It is also necessary to identify related changes in health behaviours that may be occurring at a population level in order to better understand the range of downstream psychosocial consequences of the recent outbreak. Large segments of the population under conditions of isolation need modifications to lifestyle behaviours and it is now inevitable. It changes physical activity as well as dietary habits and even the incidence of domestic violence cases. A substantial proportion of quarantined individuals during that pandemic reported being distressed and depression. A small number of researches addressed to the impacts on lifestyle or health behaviours and all are focusing on the symptoms of mental health disorders.

It is equally important to bear in mind that health behaviours are also strongly intertwined with mental health. More researches are required to the reciprocal nature of interactions between physical well-being, chronic disease and mortality, with key health behaviours such as smoking, physical activity, reduced alcohol consumption, diet and obesity. Such behaviours have an impact on mental health. Physical activity alone It has been identified as an important protective factor in reducing the risk of developing depression by Physical activity. It was reported that physical activity, alcohol consumption, smoking, body mass index and regularity of social interaction were all associated with specific mental health outcomes (i.e. depression, anxiety and stress). It is important to recognise the interactional nature of human behaviours during and after COVID-19. Some diseases like Diabetes, chronic obstructive pulmonary disease, and hypertension were the most impacted conditions due to reduction proper care. Routine care continues in spite of the pandemic for avoiding a rise in non-COVID-19-related morbidity and mortality. It is also important that patients with chronic diseases continue to receive care in spite of the pandemic [17].

III. MATERIALS AND METHODS

3.1 Neural Network and Hyper-Parameters

ML focuses on set of algorithms which are applied on dataset in order to infer predictions. ML is an art of making machines to perform efficiently without being explicitly programmed. Supervised ML is a variation of ML learning paradigm where learning algorithm accepts set of inputs along with output labels. This means that supervised ML makes the computer to discover some rule for mapping inputs to output labels [18]. Neural networks are essentially a part of ML which is simulated by the working principle of human brain. DL has proven its superiority on solving complex problems. Use of DL techniques is beneficial because, it does not include the manual feature engineering task to be performed due to its self-adaptive nature. DL often demonstrates the involvement of neural network in order to accompany complex problem solving approach. Just like neurons present in human brain, large number of

processing elements (nodes) are present in neural network for acquiring best problem solving tactic [19-20].

Before training this neural network, some pre-stage fine-tuning of hyper-parameters is necessary to perform. The hyper-parameter contains number of layers, number of nodes, learning rate, epoch size, batch size, and drop-out rate. These values should be adjusted to help the network to learn successfully. However, choosing activation function wisely is one of the necessary tasks which can maximize the training procedure. Use of activation functions allows neural networks to learn non-linear relationship among data and to produce meaningful output signal. For predicting binary class probabilities, sigmoid activation function may be used for activating output nodes. Sigmoid [21] activation function accepts the input data and transforms it in the range of 0 to 1 and it is shown in equation (1).

$$f(x) = 1/(1 + \exp^{-x}) \quad (1)$$

Tangent hyperbolic (tanh) [22] is another non-linear activation function. This is a smoother and zero-centered function. The range of this function range lies between -1 to 1, thus the output of the Tanh function is given as equation (2).

$$f(x) = (e^x - e^{-x}) / (e^x + e^{-x}) \quad (2)$$

Dropout technique is employed for case of eliminating over-fitting problem in neural networks. During training, it randomly detaches units along with incoming and outgoing connections from the neural network. Use of dropout makes neural network to acquire benchmark results in supervised classification tasks [23].

Epoch and batch size are hyper-parameters which are also used in neural network training. Both of these hyper-parameters receive integer values which need to be chosen wisely to make best use of the model's performance. Epoch size is defined to be number of passes to complete through training dataset. The entire dataset is passed forward and backward through the neural network exactly one time within each epoch. While passing the entire dataset into the algorithm, it must be partitioned into fixed size batches. Batch size keeps track of number of processed instances before the model updates its internal parameters. It is to be noted that, batch size should not be too small or too large. Having too small batch size will present high variance since small batch size may not be a good representation of the entire dataset. Again, large batch size may not be feasible because it may not fit in memory of the compute samples used for training and may lead to over-fitting problem [24].

While stacking RNN based layers into a single framework, employing an optimizer is necessary. Adam is a popular optimizer that is computationally efficient with lower memory requirement and also easy to implement. This algorithm is applicable for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order

moments. This algorithm is quite well accepted due to its applicability on non-stationary objectives and problems with very noisy and/or sparse gradients [25].

3.2 Recurrent Neural Network and GRU

Recurrent Neural Network (RNN) is constructed using multiple neural networks which are specialized for analyzing sequential data. In this network, the output of previous step is fed into current step input. For example, the output obtained during step S_i affects the parameters of step S_{i+1} . Hence, it is quite vibrant that RNN accepts two types of input- one is present input and other one is previous output for acquiring the final output. Loops present in RNN allow the signals to travel both forward and backward. Another RNN called Bidirectional RNN (BRNN) are designed to access input sequences whose starts and ends are known in advance. Because RNN can only take information from the previous context, further improvements can be made using Bi-RNN. The Bi-RNN can handle two sources of information. While considering both past and future context of each sequence element into justification, one RNN processes the sequence from start to end, the other backwards from end to start. However, RNN suffers from the problem of vanishing gradient problem.

To resolve this mentioned problem, variant of RNN Gated Recurrent Units (GRU) is explored. This variant contains gates which are neural networks that control the flow of information through the sequence chain. To mitigate short-term memory, methods like GRU introduce the concept of gates. In sequence learning tasks and overcome the problems of vanishing and explosion of gradients, Gated Recurrent unit (GRU) networks perform well when it is learning long-term dependencies. GRU consists of two types of gates such as update and reset gate. Addition and elimination of information is decided by update gate. The use of reset gate identifies how much information to hold from past. GRU uses update gate and reset gate for solving the vanishing gradient problem of a standard RNN. These vectors decide what information should be passed to the output. Without washing it through time or remove information, vectors can be trained to keep information from long ago. It is irrelevant to the prediction [9].

GRU is a gating mechanism in RNN similar to a long short-term memory (LSTM) unit. It is used without an output gate. GRU is considered a variation of the LSTM because both have a similar design and produce equal results in some cases. In GRU the update gate controls information that flows into memory and the reset gate controls the information that flows out of memory. The two vectors decide which information will get passed on to the output. GRU can be trained to keep information from the past or remove information that is irrelevant to the prediction. Given $x_t = (x_1, \dots, x_T)$ be an input sequence, W is the weight matrices σ states the sigmoid function for a GRU.

At time t , the activation function of GRU is h_t^j which is dependent on previous activation h_{t-1}^j candidate activation function h_t^j . This is formulated in equation

(3). The update gate (u_t^j), and reset gate (r_t^j) can be formulated as equation (4) and (5) respectively.

$$h_t^j = (1 - u_t^j)h_{t-1}^j + u_t^j h_t^j \quad (3)$$

$$u_t^j = \sigma(W_u \cdot [h_{t-1}^j, x_t]) \quad (4)$$

$$r_t^j = \sigma(W_r \cdot [h_{t-1}^j, x_t]) \quad (5)$$

3.3 Model Evaluation

Evaluation metrics are taken into consideration while discriminating the performance of any model from other models. Accuracy and loss are required to calculate for any deep model. For each epoch, accuracy and loss are calculated during training efficiency assessment. A loss function (or cost function) [26] measures how much the model makes mistakes for each instance in the training set. In other words, the loss function acquires the probabilities of how much predicted values get varied from original value. Accuracy and loss function values obtained for each epoch of neural network training is often plotted through graphs. Cross-entropy function can be used as loss function basically for binary classification problems. This function measures the performance of a classification model whose output is a probability value between 0 and 1 [26].

Mean Squared Error (MSE) [27] is another evaluating metric that measures absolute differences between the prediction and actual observation of the test samples. MSE produces non-negative floating point value and a value close to 0.0 turns out to be the best one. It is formulated as equation (6).

$$MSE = \frac{\sum_{i=1}^N (X_i - X_i')^2}{N}; \text{ where } X_i \text{ is the actual value and } X_i' \text{ is the predicted value} \quad (6)$$

Using true positive (TP), true negative (TN), false positive (FP), false negative (FN), accuracy and f1-score metrics can be evaluated as equation (7) and (10) respectively. It is to be noted that, f1-score is a metric that relies on calculation of recall and precision which are formulated as equation (8) and (9) respectively [27].

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+TP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (9)$$

$$F1\text{- Measure or } F1\text{-Score} = \frac{2 * \text{Recall} * \text{Precision}}{(\text{Recall} + \text{Precision})} \quad (10)$$

IV. METHODOLOGY

In order to design an automated tool for positive or negative test report prediction based on past habits and health records, a DL based framework is approached. The predictive model proposed in this paper stacks multiple GRU layers under a single platform. The GRU layers are implemented as Bi-directional RNN to possess information in both directions. These layers are followed by dropout layers. After four Bi-GRU layers and four dropout layers, four more fully connected layers are also stacked into this model. This model has been implemented using keras package [28]. The detailed

structure of this Stacked-Bi-GRU model is shown in Table 1. This model receives total 262,105 numbers of parameters and trains those parameters for obtaining

predictions. The number of layers, units present in each layer and the activation function are chosen wisely to maximize the prediction efficiency.

Layer Type	Units/Rate	Parameters Received	Activation function
Bidirectional GRU	256	99840	Sigmoid
Dropout	0.2	0	None
Bidirectional GRU	128	123264	Sigmoid
Dropout	0.2	0	None
Bidirectional GRU	64	30912	Sigmoid
Dropout	0.2	0	None
Bidirectional GRU	32	7776	Tanh
Dropout	0.2	0	None
Dense	8	264	None
Dense	4	36	None
Dense	2	10	None
Dense	1	3	Sigmoid

Table 3: Implementation summary of Stacked-Bi-GRU model

4.1 Baseline Models

Disease infection detection has also been predicted by other baseline ML models like SVM and DT. These baseline models are summarized as follows-

The advantageous in handling classification tasks with superior generalization performance are done by SVM. The upper limit of the generalization error based on the structural of risk minimization principle is minimized by SVM. It maps input vector to a higher dimensional space by constructing a maximal separating hyper-plane. For separation of data two parallel hyper-planes are constructed on each side of the hyper-plane. This hyper-plane maximizes the distance between the two parallel hyper-planes. The maximized distance between these parallel hyper-planes are considered for getting better generalization error. While handling non-linear separation problem, kernel function plays an important role. Low dimensional input space can be transformed into a higher dimensional space using SVM kernel function. In other words, kernel function applies some complex data transformations for converting non-separable problem to separable problem.

The tree-like structure is exemplified by DT and it gains knowledge on classification. Leaf node of DT is goal variable and non-leaf nodes of DT are used as a decision node. It indicates certain test that are identified by either of the branches of that decision node. Initial visit of the classifier starts from the beginning of the root this tree until a leaf node is reached. It is useful for forecasting the goal based on some criterion by implementing and training this model.

4.2 Dataset Used

This paper collects Covid-19 pre-condition patient dataset from Kaggle [29]. The dataset consists of categorical variables which are encoded to numeric form using label-encoding scheme. This step is required since ML model needs data to be accepted in numeric form. This scheme associates a number to each unique categorical value present in the column. The dataset contains numerous attributes and 566602 records those are summarized in table 2. Among all these attributes, patient id, date of symptoms arise, date of death, admission date will not contribute to the disease classification system. This analysis is based on past habit as well as other disease occurred in patients. Hence, time of symptoms, admission time in hospital, and death time are not important in this context. The target attribute 'covid result' contains values like 'positive', 'negative', 'waiting'. The records having 'waiting' covid results are eliminated from the dataset. After eliminating these records, 499692 records are currently there in the dataset. The dataset is bifurcated into training and testing dataset with a ratio of 7:3. The exact record count of testing and training dataset is summarized in table 3. The training and testing dataset is distinguished by the presence of attribute 'covid_result'. Since, prediction needs to be retrieved from testing dataset; the target attribute is eliminated from it. Any classifier model learns from the training dataset by extracting hidden patterns and utilizes that knowledge while predicting unknown sample from testing dataset.

Attributes
Id, sex, patient type (inpatient or outpatient), entry_date, date_symptoms, date_died, intubed, pneumonia, age, pregnancy, diabetes, copd, asthma, hypertension, other diseases present, cardiovascular, obesity, renal_chronic, tobacco, contact with other covid, icu, covid result

Table 2: Attributes in Dataset

V. EXPERIMENTAL RESULTS

The implemented Stacked-Bi-GRU model proceeds through training procedure for certain epochs. During one epoch training, a batch size of 64 is fitted to the model at a time. The model is trained using epoch size 5, 10 and 15. After training, testing results for accuracy is obtained. For epoch size 5, 10, 15; testing accuracies are

acquired and compared. The comparative results are shown in Fig1. From the comparative analysis, it is clear that epoch size of 10 provides better result in terms of testing accuracy than epoch size 5. However, the model is trained for 15 epochs, but the result does not improve. Hence, epoch size 10 has been chosen as best hyper-parameter.

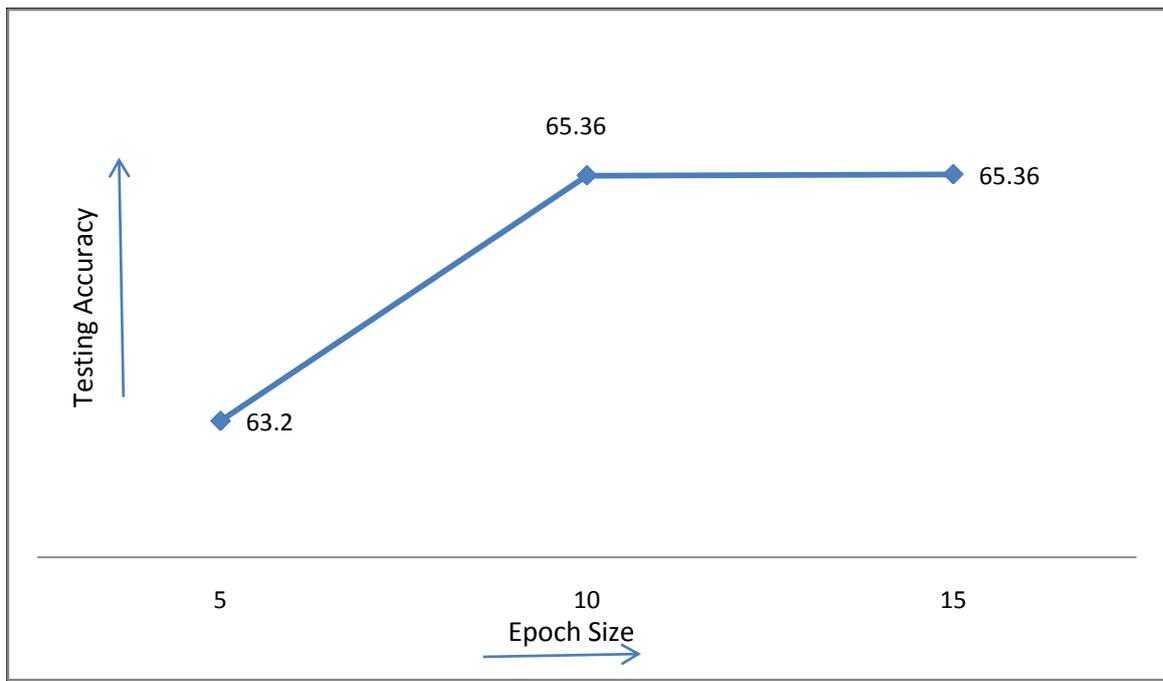


Figure 1: Testing accuracy comparison for different epoch sizes

The Stacked-Bi-GRU model having 10 epochs of training turns out to be the best model. At first epoch, the model starts with an accuracy of 0.6112 and loss of 0.6588. During 10th epoch, an accuracy of 0.6544 and loss of 0.6277 is reached. As the number of epochs increases, the model is boosted up with its accuracy while decreasing the loss. The accuracy and loss acquired for

each epoch is graphically shown in fig2. The proposed Stacked-Bi-GRU model is compared against with traditional ML classifiers such as DT and SVM. Table 3 finally summarizes the comparison among proposed classifier, DT and SVM classifier. This shows the efficiency of implemented model which is superior to specified models.

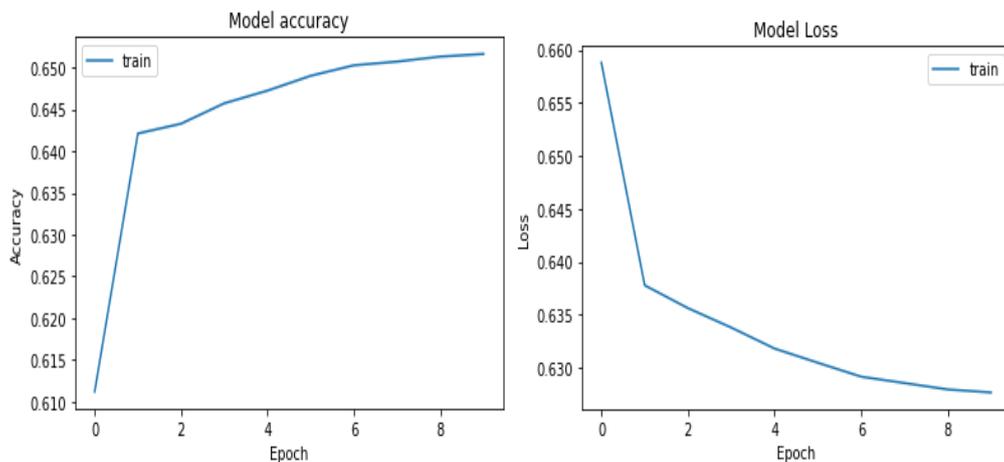


Figure 2: Training accuracy and loss acquired for 10 epochs.

Stacked Bi-GRU model	Accuracy	MSE	F1-Score
	65.36%	0.35	0.65
Decision Tree	57.32%	0.43	0.57
SVM	63.4%	0.37	0.63

Table3: Comparing Prediction Efficiency of Stacked-GRU model, SVM, DT

VI. CONCLUSION

Healthcare industry requires real time collection and processing of medical data. The main point of this industries lies the problem of data handling in real time for prediction and quick attention. DL based framework is exemplified in this paper for identifying covid-19 occurrence based on influential parameters from former habits as well as health records. This study detects the feasibility of using GRU, a gated recurrent neural network for the purpose of covid-19 disease classification. This automated tool provides comparable accuracy to the most widely used SVM, DT classifier models. This enhanced accuracy of GRU can be highly beneficial for healthcare industry.

REFERENCES

- [1] World Health Organization. *WHO Statement Regarding Cluster of Pneumonia Cases in Wuhan, China* [cited 18/08/2020]. Available at: <https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china>.
- [2] Cohen, S., Tyrrell, D. A., Russell, M. A., Jarvis, M. J., & Smith, A. P. (1993). Smoking, alcohol consumption, and susceptibility to the common cold. *American Journal of Public Health*, 83(9), 1277–1283. <https://doi.org/10.2105/ajph.83.9.1277>
- [3] Finlay Ilora, Gilmore Ian. (2020). *Covid-19 and alcohol—A dangerous cocktail*. Available at: <https://www.bmj.com/content/369/bmj.m1987>.
- [4] Centers for Disease Control and Prevention. *Symptoms of Coronavirus 2020*. [cited 18/08/2020]. Available at: <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>.
- [5] World Health Organization. (2020). *Coronavirus*. Available at: <https://www.who.int/health-topics/coronavirus>. Accessed 17 Mar 2020.
- [6] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. (2020 Feb). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 15, 395(10223), 497-506.
- [7] Marbán, Ó., Mariscal, G., & Segovia, J. (2012). *A Data Mining & Knowledge Discovery Process Model*. Available at: https://www.intechopen.com/books/data_mining_and_knowledge_discovery_in_real_life_applications/a_data_mining_amp_knowledge_discovery_process_model.
- [8] Sherstinsky A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physical D: Nonlinear Phenomena*, 404, 132306.
- [9] J. Chung, C. Gulcehre, K. Cho, & Y. Bengio. (2014). *Empirical evaluation of gated recurrent neural networks on sequence modeling*. Available at: <https://arxiv.org/abs/1412.3555>.
- [10] Yao, Y. (2001). Research on decision tree in data mining. *Application Research of Computers*.
- [11] Yang, Y., Li, J., & Yang, Y. (2015). The research of the fast SVM classifier method. In: *12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pp. 121-124.
- [12] Biddle, SJH, Garcia Bengoechea, E, & Pedisic, Z, et al. (2017) Screen time, other sedentary behaviours, and obesity risk in adults: A review of reviews. *Current Obesity Reports*, 6(2), 134–147.
- [13] Burton, R & Sheron, N. (2018). No level of alcohol consumption improves health. *The Lancet*392(10152), 987–988.
- [14] Chua, SE, Cheung, V, & Cheung, C, et al. (2004) Psychological effects of the SARS outbreak in Hong Kong on high-risk health care workers. *The Canadian Journal of Psychiatry* 49(6), 391–393.
- [15] Diener, E, Pressman, SD, & Hunter, J, et al. (2017). If, why, and when subjective well-being influences health, and future needed research. *Applied Psychology: Health and Well-Being*, 9(2), 133–167.
- [16] Dyal, SR & Valente, TW. (2015). A systematic review of loneliness and smoking: Small effects, big implications. *Substance Use & Misuse*, 50(13), 1697–1716.
- [17] Nikpouraghdam M, Farahani AJ, Alishiri G, Heydari S, Ebrahimnia M, & Samadinia H, et al. (2020). *Epidemiological characteristics of Coronavirus disease2019 (COVID-19) patients in Iran: A single Center study*. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7172806/>.
- [18] Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.
- [19] Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221–248. Available at: <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- [20] Vargas, R., Mosavi, A., & Ruiz, R. (2018). *Deep learning: A review*. Available at: <https://www.preprints.org/manuscript/201810.0218/v1>.
- [21] Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *ArXiv*, abs/1811.03378.

- [22] Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from over fitting. *J. Mach. Learn. Res.*, 15, 1929-1958.
- [23] Radiuk, P. (2017). Impact of training set batch size on the performance of convolutional neural networks for diverse datasets. *Information Technology and Management Science*, 20, 20-24.
- [24] Kingma, D.P. & Ba, J. (2015). Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980.
- [25] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. In: *NIPS Workshop on Deep Learning*.
- [26] Janocha, K. & Czarnecki, W. (2017). On loss functions for deep neural networks in classification. *ArXiv*, abs/1702.05659.
- [27] Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics*, 16(5), 412-424.
- [28] Chollet, F., & others. (2015). *Keras*. GitHub. Available at: <https://github.com/fchollet/keras>.
- [29] Tanmoy Mukherjee. (2020, Jul). *COVID-19 patient pre-condition dataset*. Version 2. Retrieved on August 8, 2020 from: <https://www.kaggle.com/tanmoyx/covid19-patient-precondition-dataset>.