Hospital Readmission Prediction Of ICU Patients Using Deep Learning Algorithms

Progress report

M.Tech Thesis Evaluation-1

November 2018

by

Shivam Sinha

(2014IPG-082)

under the supervision of

Dr. Ritu Tiwari

विश्वजीवनामृतं ज्ञानम्

ABV-INDIAN INSTITUTE OF INFORMATION TECHNOLOGY AND MANAGEMENT GWALIOR-474 010

CANDIDATE'S DECLARATION

I hereby certify that I have properly checked and verified all the items as prescribed in the check-list and ensure that my thesis/report is in proper format as specified in the guideline for thesis preparation.

I also declare that the work containing in this report is my own work. I, understand that plagiarism is defined as any one or combination of the following:

- To steal and pass off (the ideas or words of another) as one's own
- To use (another's production) without crediting the source
- To commit literary theft
- To present as new and original an idea or product derived from an existing source.

I understand that plagiarism involves an intentional act by the plagiarist of using someone else's work/ideas completely/partially and claiming authorship/originality of the work/ideas. Verbatim copy as well as close resemblance to some else's work constitute plagiarism.

I have given due credit to the original authors/sources for all the words, ideas, diagrams, graphics, computer programmes, experiments, results, websites, that are not my original contribution. I have used quotation marks to identify verbatim sentences and given credit to the original authors/sources.

I affirm that no portion of my work is plagiarized, and the experiments and results reported in the report/dissertation/thesis are not manipulated. In the event of a complaint of plagiarism and the manipulation of the experiments and results, I shall be fully responsible and answerable. My faculty supervisor(s) will not be responsible for the same.

Signature:

Name:
Roll No.:
Date:

ABSTRACT

Deep learning healthcare applications have evolved over the last few years. This advancement leads to new applications and possibilities in the field of health care. Due to different variants of deep learning algorithms like convolutional and recurrent neural network these advancements in healthcare application made possible. Healthcare has shown promises, as there is a huge share of the budget in intensive care medicine. This has expanded the interests of healthcare researchers, and they are focusing on providing the best medication to critically-ill patients. DNN models are advanced machine learning models, which have been widely applied for clinical applications in recent times. Due to the lack of ability to capture time dependency in the data which is a major shortcoming, basic machine learning algorithms like Logistic Regression, SVM, Decision Tree, etc. are not very useful. There are several types of deep learning algorithm like Deep belief networks, Deep convolutional neural network, Recurrent neural network, etc. The LSTM variant of the Recurrent neural network have shown promises to work better on data involving time series prediction, and it can transfer the results of previous cell state onto the next cell state so that dependency on time is maintained throughout. In this thesis, the proposed methodology is to use LSTM, and its different variants like Gated Recurrent Unit (GRU), LSTM attention on healthcare applications like hospital readmission prediction using ICU data. The different architectures of LSTM will produce different results based on the data on which architecture is being trained. Then the comparison of performances can be done between LSTM, and it's variants on various healthcare applications and provide the best model suitable for each of the different healthcare data being tested. The comparison of time independent machine learning models like Logistic Regression, SVM, Naive Bayes can be done with the results of Deep learning models.

Keywords: Deep learning, LSTM, Healthcare, Recurrent Neural Network, GRU.

Contents

1 Introduction

1.1 Hospital Readmission

Hospital Readmissions are identified as a signal of the poor quality of care, such as insufficient discharge planning and care coordination. Hospital readmission[28] happens when a patient discharge from hospital is readmitted within a specified interval. Readmission can occur for any planned or unplanned reason. These factors pushed Hospital organizations to think about controlling readmission rates an increase in readmission rates leads to an increase in penalties. There has been an increased usage of Readmission rates as an outcome measure in health services research and as a quality benchmark for health systems. For Medicare patients, hospitalizations are very stressful, and it is even more when they result in subsequent readmissions. A lot of research studies proved that hospitals could engage in several activities like clarifying patient discharge instructions, coordinating with post-acute care providers, vibrant cleaning mechanism to reduce the rate of readmissions of patients. But these individual follow-ups can be costly.

1.2 Machine Learning For Health Care

Machine learning leads to provide intelligence by providing new knowledge. Since the onset of machine learning, it has been widely used in achieving paramount goals as the healthcare sector is evolving. The healthcare sector is producing an enormous mass of data, and it is beyond human capability to manually analyze this vast data and produce inferences based on that. Machine learning has proven to be a boon at this level. The machine learning algorithms automatically find patterns in the data, which enable to get inferences from the data easily. Machine learning can assist healthcare providers in a variety of tasks like hospital readmission of patients suffering from chronic illness. This can not only lower the lifetime risk of the patient but also can save millions of dollars if the algorithm can predict efficiently and correctly. Other applications of machine learning in the healthcare domain includes drugs combination which should be avoided taking together, classifying images for different diseases like skin cancer. Keeping this in view, it can be said that machine learning can revolutionize the healthcare sector.

1.3 Deep Learning

Deep learning is one of the extended branches of Machine learning, which works on a specific machine learning algorithm, that is the artificial neural network. It's leveraging the recent advancements in computing power and thus using specific purpose neural network to learn from a plethora of data and make predictions based on the patterns detected. The structure consists of several hidden layers which are meant to extract the information by exploiting the structures present in the data. Deep neural networks are specialized at solving problems including data which is greatly structured. It also promises to replace hand-engineered features with feature extraction techniques that are hierarchical, unsupervised or semi-supervised. Multiple hidden units are being used between input and output in case of Deep Neural Network (DNN). Similar to artificial neural network, the complex non-linear relationships can be modeled using DNN. .[2]

Similar to artificial neural network, the complex non-linear relationships can be modeled using DNN.

1.4 Recurrent Neural Network

An artificial neural network has a special type called recurrent neural network (RNN), and it has directed cycles between connection. The dynamic temporal behavior is exhibited by the internal state of the network. Unlike ANN, arbitrary sequences can be processed by using the internal state of RNN. This is the fundamental architecture developed in the 1980s: neuron-like units in a network, a directed connection between every neuron. There is a time-varying real-valued activation in every unit. There is modifiable real-valued weight in each connection. There are input, output, and hidden nodes. One input vector is supplied at every time steps. At every time step, the non-linear function of the weighted sum of all the activations of connected units is calculated by each non-input unit.[23].

1.5 Long Short Term Memory

The Long short-term memory (LSTM) network, a deep learning RNN, is used by numerous researchers, published by Schmidhuber & Hochreiter in 1997[19]. This deep learning system overcomes the vanishing gradient problem of traditional RNN. The recurrent gates called forget gates are augmented in LSTM. The vanishing or exploding of back-propagated error is prevented in LSTM RNNs. There are an unlimited number of virtual layers in LSTM RNNs, through which the error can flow backward. LSTM can learn from events that occurred thousands of time steps ago so it can be used for very deep learning tasks. LSTM has a unique architecture that consists of a memory cell, which can maintain the state which it is currently in, over time. Non- Linear gating units are used in LSTM to regulate the information flow in and out of the memory cell. Over time, many improvements have been suggested and made in the standard architecture of LSTM to make it more efficient. Today, LSTMs are being used in a variety of learning problems which differ in scale and nature significantly when compared to the problems which were initially solved by LSTM.

2 Review of key related research

2.1 Background

A procedure of the current work was performed with a specific end goal to study about the most significant heathcare applications literature for the survey and also to study an expansive piece of the most profound learning approaches for healthcare applications.

2.2 Related Research and Analysis

Author (Year)	**Meta-heuristic**	**Paper Name**	**Comments**
Ghassemi et al. (2014) [14]	Multivariate Logistic Regression	A data-driven approach to optimized medication dosing: a focus on heparin	An approach is developed that help clinicians determine the optimal initial dose of a drug to safely and quickly reach a therapeutic aPTT window.
Pirracchio et al. (2015)[24]	Super ICU Learner Algorithm	Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study.	A super learner algorithm is used for predicting hospital mortality in patients.
Che et al. (2015) [6]	Stacked Denoising Autoencoder & Long Short-Term Memory	Distilling knowledge from deep networks with applications to healthcare domain	A novel called Interpretable Mimic Learning is used, to learn interpretable phenotype features for making robust prediction while mimicking the performance of deep learning models meta-heuristics.
Liang and Hu (2015) [22]	Convolution & Recurrent Neural Network	Recurrent convolutional neural network for object recognition	Combination of convolution and recurrent neural network model is used for object detection.
Ioffe and Szegedy (2015) [20]	Batch Normalization	Batch normalization: Accelerating deep network training by reducing internal covariate shift	Reduction in overfitting using batch normalization.

Author (Year)	Meta-heuristic	Paper Name	Comments
Che et al. (2016) [7]	Knowledge Distillation Approach	Interpretable deep models for icu outcome prediction	Interpretable mimic learning is introduced that uses gradient boosting trees to learn interpretable models and at the same time achieves strong prediction performance as deep learning models.
Choi, Bahadori, Schuetz, Stewart and Sun (2016) [9]	Recurrent Neural Networks	Doctor ai: Predicting clinical events via recurrent neural networks	The proportional split of data is found useful for deep learning applications when dealing with large volume of data.
Choi, Schuetz, Stewart and Sun (2016) [10]	Gated Recurrent Units	Using recurrent neural network models for early detection of heart failure onset	Leveraging longitudinal EHR data for early detection of heart failure using Recurrent Neural Network.
Hanson et al. (2016) [18]	Bidirectional LSTM	Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks	Long short term memory rnn is used to predict disordered proteins, such as SPOT-disorder.
Johnson et al. (2016) [21]	Dataset Extraction	MIMIC-III, a freely accessible critical care database	Medical Information Mart for Intensive Care (MIMIC-III) consists of data about patients admitted to various critical care units in a large hospital.
Che et al. (2017) [8]	Recurrent Neural Networks	Deep Learning Solutions for Classifying Patients on Opioid Use	State-of-the-art deep and recurrent neural network models were applied to achieve robust results on classifying opioid users.
Dey and Salemt (2017) [11]	Gated Recurrent Unit	Gate-variants of Gated Recurrent Unit (GRU) neural networks	Solve the vanishing gradient problem of a standard RNN.
Greff et al. (2017) [17]	LSTM	LSTM: A search space odyssey	Overcomes the vanishing gradient problem of traditional RNN.

Author (Year)	Meta-heuristic	Paper Name	Comments
Rubin et al. (2017) [26]	Convolution Recurrent neural Network	Recognizing abnormal heart sounds using deep learning	Combines the use of time-frequency heat map representations with a deep convolutional neural network (CNN) to classify heart sounds.
Franco et al. (2017) [13]	Statistical Algorithms	Impact of prealbumin on mortality and hospital readmission in patients with acute heart failure	Statistical Analysis is used to find the mortality and readmission rates.
Esteva et al. (2017) [12]	Convolution Neural Network	Dermatologist-level classification of skin cancer with deep neural networks	Training of Skin lesions using a single CNN is done on an end-to-end basis from images directly, using only two inputs- pixels and diseases.
Biswal et al. (2017) [3]	SLEEPNET	SLEEPNET: automated sleep staging system via deep learning	Propose SLEEPNET (Sleep EEG neural network), a deployed annotation tool for sleep staging.
Che et al. (2018) [5]	Gated Recurrent Unit	Recurrent neural networks for multivariate time series with missing values	Uses GRU-D to capture long-term temporal dependencies in time series and utilizes the missing patterns to achieve better prediction results.
Reddy and Delen (2018) [25]	LSTM	Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology	Utilizes deep learning methods to predict rehospitalization within 30 days by extracting the temporal relationships in the longitudinal EHR clinical data.
Xiong et al. (2018) [29]	Convolution Recurrent Neural Network	ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network	Developed RhythmNet, a 21-layer 1D convolutional recurrent neural network to classify ECGs of different rhythms including AF automatically.

Author (Year)	Meta-heuristic	Paper Name	Comments
Biswas et al. (2018) [4]	Convolution based encoder & decoder	Deep learning strategy for accurate carotid intima-media thickness measurement: An ultrasound study on Japanese diabetic cohort	The Deep Learning system can be used for stroke risk assessment during routine or clinical trial modes.
Acharya et al. (2018) [1]	Convolution Neural Network	Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals	A Novel method for UCAV path planning in 2D environment. A 13-layer deep learning CNN algorithm is implemented for the automated EEG analysis.

2.3 Research Gaps

In literature survey we find the following research gaps. These are points that we will be focused on in this thesis.

- The results obtained in these techniques can be improved using novel classification algorithms. Most of the researchers have used existing machine learning algorithms, and it is expected that deep learning techniques will surpass the conventional techniques.
- These previous works were not able to model high dimensional nonlinear relations as good as RNN.
- Descriptive statistics were being used in earlier methods. However, these statistics like mean, median & mode are always under the risk of losing some vital information.
- Deep learning algorithms are still under a shadow for a variety of healthcare applications.

3 Objectives

The primary objective of this thesis is:

- To develop Deep Neural Network(DNN) models for Healthcare application, i.e., Hospital Readmission.
- To predict hospital readmission of patients using deep neural networks.
- To compare results between Deep Learning Models and with existing algorithms.

4 Methodology

4.1 Data description: MIMIC-III

Medical Information Mart for Intensive Care (MIMIC-III)[21] consists of data about patients admitted to various critical care units in a large hospital. MIMIC-III is generally viewed as a large and single-center database. A large number of different parameters are present in MIMIC III database. These parameters include information such as vital signs, medications, laboratory measurements, observations, fluid balance, procedure codes, diagnostic codes, imaging reports, hospital length of stay, survival data, and others. The database consists of information of around 58,576 distinct patients who were admitted to various critical care units of the hospital between 2001 and 2012. The data comprises of patients aged 16 years or above only. Descriptive statistics were performed on this dataset, and it was found that the median age is 65.8 years for adult patients(Q1-Q3: 52:877:8). Out of total patients, there were 55.9% male patients and only 11.5% of the cases had in-hospital mortality. It was calculated from the data that on an average a patient stayed for 2.1 days in ICU and had an average of 6.9 Days of hospital stay. MIMIC-III database have several idiosyncratic properties about it, and these are mentioned below:

- The dataset is a huge database accumulated throughout more than a decade and consists of very detailed information of each patient under care.
- MIMIC-III database requires a user to oblige by a user data agreement and once itis accepted the analysis is boundless.
- It contains discharge summaries as well as reports of ECG, imaging studies and information about various codes like International Classification of Disease, 9thEdition (ICD-9) codes, Diagnosis Related Group (DRG) codes, and Current Procedural Terminology (CPT) codes.
- It is a time series data. Clinical variables are recorded concerning time for each patient. Figure 1 shows the chart of timing clinical variables recorded in the dataset.

4.2 Data Preprocessing

4.2.1 Data Extraction

MIMIC-III dataset consisted of around 58,576 patients who were diagnosed as suffering from different types of disease and hence mortality due to those diseases. These diseases include Pulmonary disease, Circulatory disease, Trauma, a disease of the digestive system and many more. Since the dataset is very large, we only consider data of those patients who were readmitted again which gives the details of 7,534 patients. The data set is divided into two classes:-

- Patients readmitted in 30 days.

- Patients readmitted after 30 days.

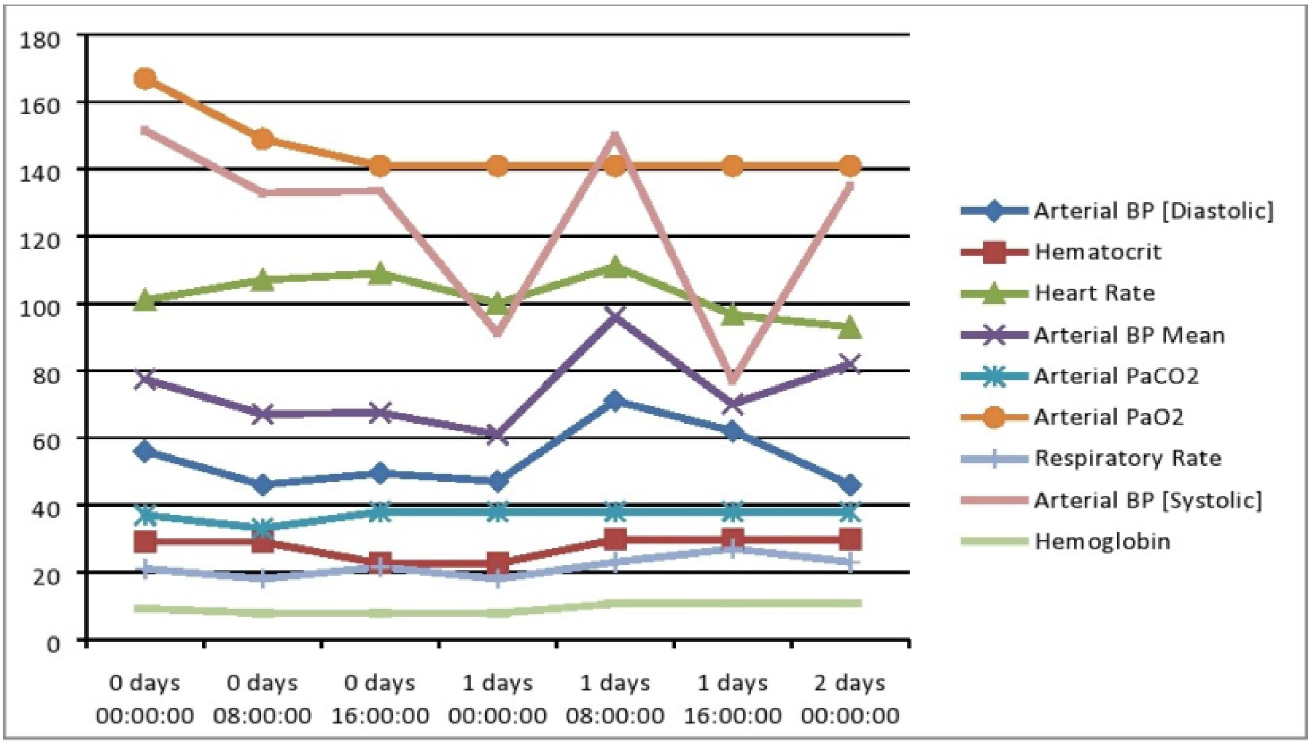

Figure 1: Some of the Clinical recordings associated with each patient[21]

4.2.2 Missing Values

MIMIC-III dataset contains missing values in some of the features. The feature will be removed if it contains more missing value otherwise mean will be used to fill the missing values.

4.2.3 Normalization

Normalization is used to remove the biases among the features. It brings the data on a standard scale. It standardizes the range of independent features or variables of data, called feature scaling.

4.3 Algorithms

4.3.1 Recurrent Neural Network

Recurrent neural networks (RNN) are used to capture the temporal dependency in the time series data[5]. As most of the healthcare data is a series of temporal recordings, hence RNNs have been widely adopted in the healthcare domain. We are usually provided with a series of observations $x_1 ... x_T$ and we train a classifier to generate hypotheses ŷ. The recurrent connections are added in feed-forward neural networks to make it RNN. The output of a neuron in a typical NN is as follows:

$$y_i^t = \sigma(W_i x^t + b_i) \tag{1}$$

Where W_i is the weight matrix, bi is the bias and represents the sigmoid function. While in the case of RNN, a neuron is fed with the output of the neuron at time t-1. The following equation shows the new activation function:

$$y_i^t = (W_i x^t + V_i x^{t-1} + b_i) \tag{2}$$

As RNN uses the previous outputs as recurrent connection, their current output depends upon the previous states. This property of RNN makes it very useful in sequence labeling tasks. The backpropagation through time can be used to train RNNs. It was demonstrated by that learning long-term dependencies is difficult using gradient descent. This is mainly because the backpropagating error can vanish which makes the network inefficient in learning long-range dependencies, or frequently explode which makes convergence impossible.

LSTM networks were proposed to tackle the problem of vanishing gradients and were developed to model long-range dependencies efficiently. LSTMs can accomplish this by keeping an internal state that represents the memory cell of the LSTM neuron. This internal state can only be written and read through gates which control the information flowing through the cell state. The following diagram shows the recurrent neural network.

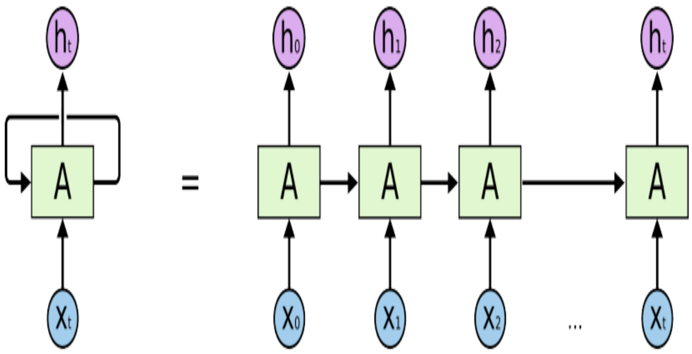

Figure 2: An unrolled recurrent neural network[17]

4.3.2 Long Short Term Memory (LSTM) Networks

Long Short Term Memory networks usually just called LSTMs are a special kind of RNN, capable of learning long-term dependencies. They were introduced by [19] and were refined and popularized by many people. They work remarkably well on a large variety of problems and are now widely used. To solve long-term dependency problem, LSTMs are explicitly designed. Remembering information for long periods of time is practically their default behavior, not something they struggle to learn! All recurrent neural networks have the form of a chain of repeating modules of the neural network. In traditional RNNs, this repeating module will have a pretty simple structure, such as a single tank layer. LSTMs also have this chain-like structure, but the recurrent module has a different structure. Rather than having a single neural network layer, there are four layers which are interacting extraordinarily. The intermediate information is stored in a single hidden layer h and its state changes over time (h_{t1}, h_t, h_{t+1}). On the final hidden state vector h_T, we used a fully connected layer followed by sigmoid function. For loss function, we used log loss. The following equations can be used for calculation of the current hidden layer h_t.

$$f_t = \sigma(W_f X_t + R_f h_{t1} + b_f) \tag{3}$$

$$i_t = \sigma(W_i X_t + R_i h_{t1} + b_i) \tag{4}$$

$$o_t = \sigma(W_o X_t + R_o h_{t1} + b_o) \tag{5}$$

$$\tilde{C}_t = \Phi(W_C X_t + R_C h_{t1} + b_c) \tag{6}$$

$$C_t = f_t C_{t1} + i_t \tilde{C}_t \tag{7}$$

$$h_t = o_t \phi(C_t) \tag{8}$$

f_t , i_t and o_t are forget gate, input gate and output gate respectively. To decide which historical information will be discarded from the cell state forget gates are used, the update of cell state is decided by the input gate, and the output gate decides the output of the cell state. Cell states are completely overridden in classical RNN, but LSTM has the potential to add or remove information to the cell state. The input weight of each gate, recurrent weight, and the bias are expressed as W* , R* , b* respectively where * can be f, i, o and c. Here σ, Φ stands for an element-wise application of the sigmoid (logistic) and tanh function respectively. For matrix multiplication. The candidate values are computed in equation (6), and equation(7) old state is multiplied by f_t, and this helps in forgetting the things we decided to forget. Then i_t*δC_t is added in it. This is the new candidate values, scaled by how much we decided to update each state value. The final output of an LSTM unit is given by equation 8. Here * represents the Hadamard (element-wise) multiplication operation. Figure 3 shows the LSTM network.

4.3.3 Bidirectional Recurrent Neural Networks (BRNN)

Bidirectional Recurrent Neural Networks also called BRNN is just like RNN but it trains simultaneously on both sides of the time series data[27]. This model gives the better

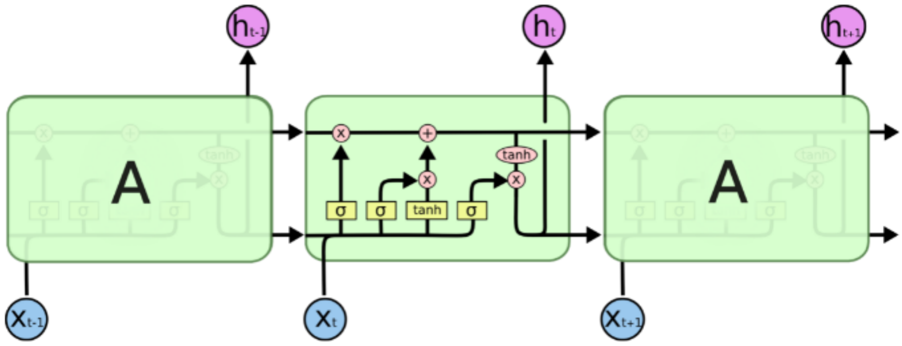

Figure 3: The repeating module in an LSTM contains four interacting layers[17]

result in both regression and classification problem. BRNN computes both forward ($\overrightarrow{\mathrm{h}}$) and backward ($\overleftarrow{\mathrm{h}}$) hidden sequence.

$$\overrightarrow{\mathrm{h}}_t = \mathcal{H}(W_{x\overrightarrow{\mathrm{h}}}x_t + W_{\overrightarrow{\mathrm{h}}\overrightarrow{\mathrm{h}}}\overrightarrow{\mathrm{h}}_t + b_{\overrightarrow{\mathrm{h}}}) \tag{9}$$

$$\overleftarrow{\mathrm{h}}_t = \mathcal{H}(W_{x\overleftarrow{\mathrm{h}}}x_t + W_{\overleftarrow{\mathrm{h}}\overleftarrow{\mathrm{h}}}\overleftarrow{\mathrm{h}}_t + b_{\overleftarrow{\mathrm{h}}}) \tag{10}$$

$$y_t = W_{\overrightarrow{\mathrm{h}}y}\overrightarrow{\mathrm{h}}_t + W_{\overleftarrow{\mathrm{h}}y}\overleftarrow{\mathrm{h}}_t + b_o \tag{11}$$

The long range context can be accessed in both directions by combining BRNNs with LSTM which gives bidirectional LSTM[16].

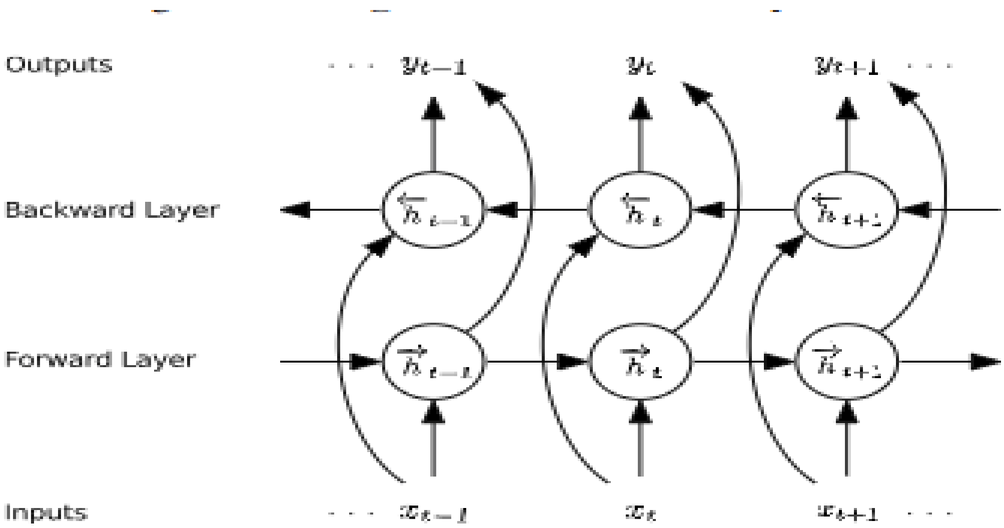

Figure 4: Bidirectional Recurrent Neural Networks[15]

4.3.4 Flow Diagram

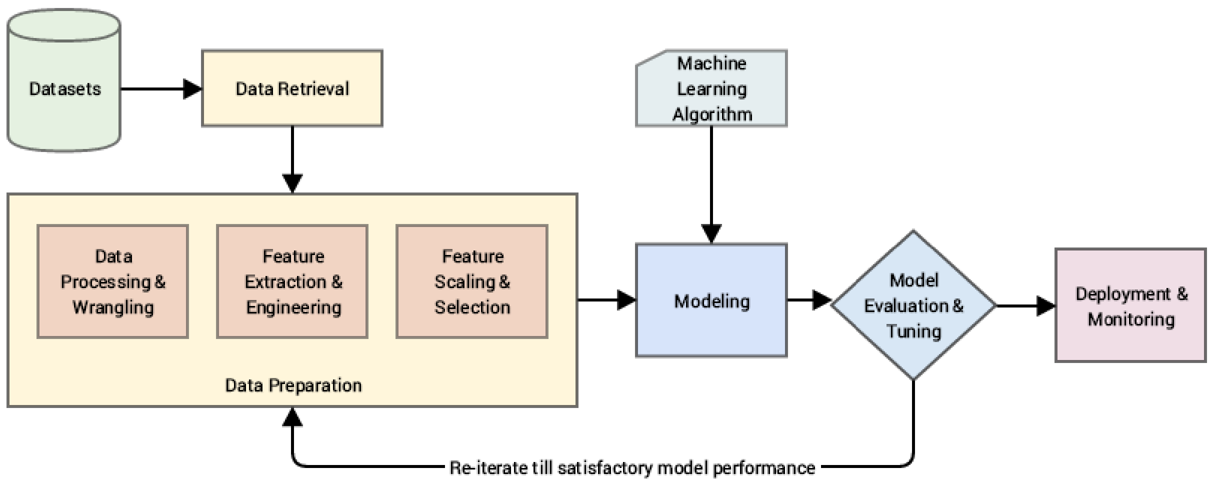

Figure 5: Flow Diagram

5 Expected Results

- Accuracy (on Test Data)
- AUC (Area Under ROC Curve).
- Enhancement of an existing algorithm(RNN, RNN-LSTM).

References

[1] Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H. and Adeli, H.: 2018, Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals, *Computers in biology and medicine* **100**, 270–278.

[2] Bengio, Y. et al.: 2009, Learning deep architectures for ai, *Foundations and trends® in Machine Learning* **2**(1), 1–127.

[3] Biswal, S., Kulas, J. A., Sun, H., Goparaju, B., Westover, M. B., Bianchi, M. T. and Sun, J.: 2017, Sleepnet: Automated sleep staging system via deep learning, *CoRR* **abs/1707.08262**.

[4] Biswas, M., Kuppili, V., Araki, T., Edla, D. R., Godia, E. C., Saba, L., Suri, H. S., Omerzu, T., Laird, J. R., Khanna, N. N. et al.: 2018, Deep learning strategy for accurate carotid intima-media thickness measurement: An ultrasound study on japanese diabetic cohort, *Computers in biology and medicine* **98**, 100–117.

[5] Che, Z., Purushotham, S., Cho, K., Sontag, D. and Liu, Y.: 2018, Recurrent neural networks for multivariate time series with missing values, *Scientific reports* **8**(1), 6085.

[6] Che, Z., Purushotham, S., Khemani, R. G. and Liu, Y.: 2015, Distilling knowledge from deep networks with applications to healthcare domain, *CoRR* **abs/1512.03542**.

[7] Che, Z., Purushotham, S., Khemani, R. and Liu, Y.: 2016, Interpretable deep models for icu outcome prediction, *AMIA Annual Symposium Proceedings*, Vol. 2016, American Medical Informatics Association, p. 371.

[8] Che, Z., Sauver, J. S., Liu, H. and Liu, Y.: 2017, Deep learning solutions for classifying patients on opioid use, *AMIA Annual Symposium Proceedings*, Vol. 2017, American Medical Informatics Association, p. 525.

[9] Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F. and Sun, J.: 2016, Doctor ai: Predicting clinical events via recurrent neural networks, *Machine Learning for Healthcare Conference*, pp. 301–318.

[10] Choi, E., Schuetz, A., Stewart, W. F. and Sun, J.: 2016, Using recurrent neural network models for early detection of heart failure onset, *Journal of the American Medical Informatics Association* **24**(2), 361–370.

[11] Dey, R. and Salemt, F. M.: 2017, Gate-variants of gated recurrent unit (gru) neural networks, *Circuits and Systems (MWSCAS), 2017 IEEE 60th International Midwest Symposium on*, IEEE, pp. 1597–1600.

[12] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. and Thrun, S.: 2017, Dermatologist-level classification of skin cancer with deep neural networks, *Nature* **542**(7639), 115.

[13] Franco, J., Formiga, F., Trullas, J.-C., Bautista, P. S., Conde, A., Manzano, L., Quirós, R., Franco, Á. G., Ezquerro, A. M., Montero-Pérez-Barquero, M. et al.: 2017, Impact of prealbumin on mortality and hospital readmission in patients with acute heart failure, *European journal of internal medicine* **43**, 36–41.

[14] Ghassemi, M. M., Richter, S. E., Eche, I. M., Chen, T. W., Danziger, J. and Celi, L. A.: 2014, A data-driven approach to optimized medication dosing: a focus on heparin, *Intensive care medicine* **40**(9), 1332–1339.

[15] Graves, A. and Jaitly, N.: 2014, Towards end-to-end speech recognition with recurrent neural networks, *International Conference on Machine Learning*, pp. 1764–1772.

[16] Graves, A. and Schmidhuber, J.: 2005, Framewise phoneme classification with bidirectional lstm and other neural network architectures, *Neural Networks* **18**(5-6), 602–610.

[17] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. and Schmidhuber, J.: 2017, Lstm: A search space odyssey, *IEEE transactions on neural networks and learning systems* **28**(10), 2222–2232.

[18] Hanson, J., Yang, Y., Paliwal, K. and Zhou, Y.: 2016, Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks, *Bioinformatics* **33**(5), 685–692.

[19] Hochreiter, S. and Schmidhuber, J.: 1997, Long short-term memory, *Neural computation* **9**(8), 1735–1780.

[20] Ioffe, S. and Szegedy, C.: 2015, Batch normalization: Accelerating deep network training by reducing internal covariate shift, *Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, JMLR.org, pp. 448–456.
URL: *http://dl.acm.org/citation.cfm?id=3045118.3045167*

[21] Johnson, A. E., Pollard, T. J., Shen, L., Li-wei, H. L., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A. and Mark, R. G.: 2016, Mimic-iii, a freely accessible critical care database, *Scientific data* **3**, 160035.

[22] Liang, M. and Hu, X.: 2015, Recurrent convolutional neural network for object recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3367–3375.

[23] Lipton, Z. C.: 2015, A critical review of recurrent neural networks for sequence learning, *CoRR* **abs/1506.00019**.

[24] Pirracchio, R., Petersen, M. L., Carone, M., Rigon, M. R., Chevret, S. and van der Laan, M. J.: 2015, Mortality prediction in intensive care units with the super icu learner algorithm (sicula): a population-based study, *The Lancet Respiratory Medicine* **3**(1), 42–52.

[25] Reddy, B. K. and Delen, D.: 2018, Predicting hospital readmission for lupus patients: An rnn-lstm-based deep-learning methodology, *Computers in biology and medicine* **101**, 199–209.

[26] Rubin, J., Abreu, R., Ganguli, A., Nelaturi, S., Matei, I. and Sricharan, K.: 2017, Recognizing abnormal heart sounds using deep learning, *KHD@IJCAI*.

[27] Schuster, M. and Paliwal, K. K.: 1997, Bidirectional recurrent neural networks, *IEEE Transactions on Signal Processing* **45**(11), 2673–2681.

[28] Strack, B., DeShazo, J. P., Gennings, C., Olmo, J. L., Ventura, S., Cios, K. J. and Clore, J. N.: 2014, Impact of hba1c measurement on hospital readmission rates: analysis of 70,000 clinical database patient records, *BioMed research international* **2014**.

[29] Xiong, Z., Nash, M. P., Cheng, E., Fedorov, V. V., Stiles, M. K. and Zhao, J.: 2018, Ecg signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network, *Physiological measurement* **39**(9), 094006.