

# Prediction of Stroke using Deep Learning Model

Pattanapong Chantamit-o-pas<sup>1</sup> and Madhu Goyal<sup>2</sup>

<sup>1,2</sup>Centre for Artificial Intelligence, Faculty of Engineering and Information Technology  
University of Technology Sydney, PO BOX 123, Broadway, NSW 2007, Australia

Pattanapong.Chantamit-o-pas@student.uts.edu.au  
Madhu.Goyal-2@uts.edu.au

**Abstract.** Many predictive techniques have been widely applied in clinical decision making such as predicting occurrence of a disease or diagnosis, evaluating prognosis or outcome of diseases and assisting clinicians to recommend treatment of diseases. However, the conventional predictive models or techniques are still not effective enough in capturing the underlying knowledge because it is incapable of simulating the complexity on feature representation of the medical problem domains. This research reports predictive analytical techniques for stroke using deep learning model applied on heart disease dataset. The atrial fibrillation symptoms in heart patients are a major risk factor of stroke and share common variables to predict stroke. The outcomes of this research are more accurate than medical scoring systems currently in use for warning heart patients if they are likely to develop stroke.

**Keywords:** Deep learning, Predictive techniques, Stroke.

## 1 Introduction

Data mining has become an essential instrument for researchers and clinical practitioners in medicine and numerous reports using these techniques are available. However, one of the biggest problems in data mining in medicine is that medical data is voluminous, heterogeneous and complex. The need for algorithms with very high accuracy is required as medical diagnosis is considered quite significant task that needs to be carried out precisely and efficiently.

Most common techniques that used to induce predictive models from data set are naïve Bayesian classifier and the decision tree. Bayesian classifier is one of the simplest yet a fairly accurate predictive data mining method [1]. A clinical decision support system [2] is used for prediction and diagnosis in heart disease. This approach is able to extract hidden pattern and relationships among medical data for prediction of heart disease using major risk factors. It applied genetic algorithms and neural networks and is called 'hybrid system'. It used a genetic algorithm feature for initialization of neural network weights.

For stroke, the predictive techniques range from simple models to more complex predictive models. Risk factors of stroke are complex and applied for finding different

complexities of disease and uncertainty from direct and/or non-direct risk factor sources. The analysis of stroke patients who were admitted in the TOAST study was done using stepwise regression methods [3]. This research used 1,266 stroke patients from database who had suffered in a transient ischemic attack (TIA) or recurrent stroke within 3 months after stroke, and selection 20 clinical variable for finding performance and evaluation.

The prognostic significance of blood pressure for stroke risk was examined by using the Cox proportional hazards regression model, which was adjusted for possible confounding factors. The results from a number of measurements showed that the predictive value of home blood pressure increased progressively. The initial home blood pressure values (one measurement) showed a significantly greater relation with stroke risk than conventional blood pressure values [4].

The Cox proportional hazards model and machine learning approach have been compared for stroke prediction on the Cardiovascular Health Study (CHS) dataset [5]. Specially, they considered the common problems of prediction in medical dataset, feature selection, and data imputation. This research proposes the use of an innovative algorithm for automatic feature selection - which chooses robust features based on heuristic: conservative mean. This algorithm was applied in combination with Support Vector Machines (SVMs). The feature selection algorithm achieves a greater area under the ROC curve (AUC) in comparison with the Cox proportional hazards model and L1 regularized Cox model. The method was also applied to clinical prediction of other diseases - where missing data are common, and risk factors are not well understood.

The Bayesian Rule Lists generated stroke prediction model employing the Market Scan Medicaid Multi-State Database (MDCD) with Atrial Fibrillation (AF) symptom[6]. The database categorised 12,586 patients on the basis of AF diagnosis. The observation was divided into two phases: a 1-year observation prior to the diagnosis; and 1-year observation after the diagnosis. The result found that 1,786 patients had a stroke within a year after suffering the atrial fibrillation. With regards to evaluation, the Bayesian Rule List (BRL) point is estimated by constructing a receiver operating characteristic (ROC) curve and measuring area under the curve (AUC) for each fold.

The classification algorithms i.e. Neural Network, Naïve Bayes, and Decision Tree are used for predicting the presence of stroke with various related attributes. The principle component analysis algorithm is commonly used for reducing the dimensions. Also, it is used for determining more relevant attributes towards the prediction of stroke and predicting whether the patient is suffering from stroke or not [7].

The conventional models are incapable of detecting fundamental knowledge because they fail to simulate the complexity and feature representation of medical problem domains. Researchers attempt to apply a deep model to overcome this weakness. Several applications of deep learning model to medical data analysis have been reported in recent years, for instance, an image analysis system for histopathological diagnosis on the images. Liang, Zhang, Huang and Hu [8] suggested the application of deep belief network for unsupervised feature extraction, and then perform supervised learning through a standard SVM. The results confirm the advantage of deep

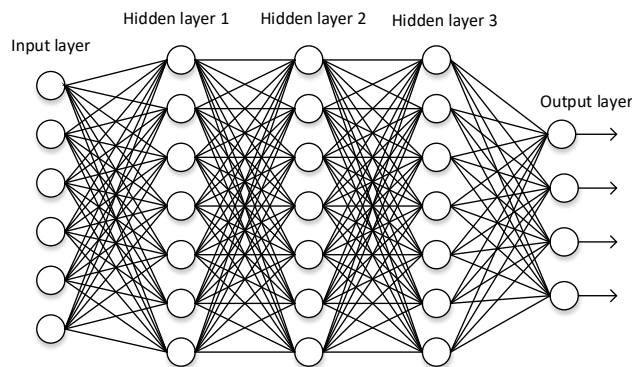
model towards knowledge modelling for data from medical information systems such as Electronic Medical Record (EMR) and Hospital Information System (HIS). Thus, predictive analytical techniques for stroke using deep learning techniques are potentially significant and beneficial.

The rest of this paper is organized as follows. Section 2 reviews deep learning technique in healthcare sector. Section 3 discusses the back propagation algorithm. Section 4 discusses application of deep learning model on heart disease dataset and the conclusion and future work are presented in section 5 of this paper.

## 2 Deep Learning

Recently researchers [8-10] have been using deep learning technique for prediction. This technique employs learning from data with multiple level of abstraction by computational models that are associated with multiple processing layers. This method intended to discover complex structure in big data set by using the back propagation algorithm to predict the result. The machine can learn from source and change its internal parameters by computing the representation in each layer to form the representation in the previous layer [11].

Deep Learning architecture can be illustrated in form of different concept levels that uncover hidden layers in a problem domain. Deep learning models represent dataset in a multi-layer form. Each layer derives from the computation of node and weight of connections among nodes, and each transform represents one level, which will be the input for the next layer (see Fig 1).



**Fig. 1.** A structure of deep neural network

In healthcare studies, researchers [10] developed health seekers by searching information needed for predicting possible disease, and there were able to show an evidence of symptoms for further analysis. Their research proposed the use of deep learning technique to find the possible disease based on the question of health seekers. The main discussion of research included two keys component. The first component is the detachment of the discriminant medical signatures form raw features. The next component is the input of raw features and their signatures into the first layers and

hidden nodes. Meanwhile, the inter-relations in each layer learned between pre-training and pseudo-labelled data. In hidden layers, the raw features served the abstract signature mining. Therefore, the repetition and increase of these two components build a deep architecture to connect with three hidden layers.

Basically, a stroke is where an area of brain gets deprived of its blood supply, and it can happen by a vessel being blocked, which is then called an “*ischemic stroke*”. Three quarters of ischemic stroke result in a case of “*hemorrhagic stroke*” where a blood vessel can burst.

In acute ischemic stroke treatment, the prediction of tissue survival outcome plays a fundamental role in the clinical decision-making process as it can be used to assess the balance of risk and possible benefit when considering endovascular clot retrieval intervention. For the first time, Stier, Vincent, Liebeskind and Scalzo [9] constructed a deep learning model of tissue fate based on randomly sampled local patches from the hypoperfusion (Tmax) feature observed in MRI immediately after symptom onset. They evaluated the model with respect to the ground truth established by an expert neurologist four days after intervention. The results show the superiority of the proposed regional learning framework versus a single-voxel-based regression model. The previous research reveal to kernel of the deep learning techniques can be applied to healthcare sector as regulariser at the output layers or a part of model.

### 3 Deep learning Architecture for Heart Disease Dataset

Features selection and classification analysis can be done by algorithms such as, Naïve Bayes, Decision Tree, SVM and many other techniques. These techniques extract knowledge from large databases for crucial decision support. A supervised learning technique is applied to healthcare applications for supporting medical diagnosis, improve patient care, and decision-making treatment [12]. This technique relies on a training dataset when constructing a prediction model for a target task. Each training is described various variables such as weights, classes, and attributes. The training set will compare between previous cases and new cases. However, diagnosis of disease is difficult problem because the number of risk factors are increase and complex. So, it need to improve prediction accuracy.

Backpropagation algorithm (BP) is applied to the main deep model as for learning in multi-layer networks in this study. This model is supervised concept extractor for the original dataset samples. A backpropagation is treated as a multi-layer feedforward neural networks with hidden layers or multiple-layer neural network (see Fig 1). Each hidden unit can be considered as multiple outputs perceptron network, so we can be considered a soft-threshold linear combination of the hidden units, which are equivalent to the output unit perceptron.

We need to consider multiple output unit for multi-layer networks. Let  $(x, y)$  be a single sample with its desired output labels  $y = \{y_1, \dots, \dots, y_i\}$ . The error at the output unit is just  $y - h_w(x)$ , and it can use this to adjust the weights between the hid-

den layers and the output layers, so this process that produce the error at the hidden layers in terms of equivalent to the error at the hidden layers. This is subsequently used for update the weights between the input units and the hidden layers as in algorithm (see Fig 2).

```

function BP-LEARNING returns a neural network
inputs: examples, a set of data, each with input vector  $\mathbf{x}$  and output vector  $\mathbf{y}$  network, a multi-layer network with L layers, weights  $W_{j,i}$ , activation function  $g$ 
repeat
  for each  $e$  in examples do
    for each node  $j$  in the input layer do  $a_j \leftarrow x_j[e]$ 
    for  $\ell = 2$  to  $M$  do
       $in_i \leftarrow \sum_i W_{j,i} a_j$ 
       $a_i \leftarrow g(in_i)$ 
    for each node  $i$  in the output layer do
       $\Delta_j \leftarrow g'(in_j) \times (y_i[e] - a_i)$ 
    for  $\ell = M - 1$  to  $1$  do
      for each node  $j$  in layer  $\ell$  do
         $\Delta_j \leftarrow g'(in_j) \sum_i W_{j,i} \Delta_i$ 
        for each node  $i$  in layer  $\ell + 1$  do
           $W_{j,i} \leftarrow W_{j,i} + \alpha \times a_j \times \Delta_i$ 
    until some stopping criterion is satisfied
return NEURAL NETWORK

```

**Fig. 2.** The backpropagation algorithm for learning in multi-layer networks [13].

After different feature representations have been introduced to the network, the weight parameters automatically tune themselves to a supervisee feature. Augments of the previous steps become input to the next layer. The larger layers are, the faster model can compute. Either a classification or regression model is required for the hidden layers, and by virtue of it, the hidden layers can serve in filtering final output. However, it is required to decide again whether the outputs are valid.

## 4 Model evaluation

### 4.1 Data source

This research has applied deep learning model on heart disease dataset (available at UCI Machine learning website). It has 899 records and 76 attributes per record. It contains Patient Number, Social Security Number, Age, Gender, Blood pressure, type of chest pain, Cigarettes, Family history, Hypertension, Cholesterols, Years, EKG (day/month/year), Heart rate, Nitrates, and calcium channel blocker, and so on. It covers four hospitals at medical center in Hungarian, Switzerland, Cleveland and Long Beach Virginia. From this dataset, we selected ten attributes such as Age, Gender, Blood Pressure, Chest Pain, Cigarettes, Family History, Hypertension, Cholester-

ols, Heart Rate, and blood vessels. These attributes are related to stroke's risk factor used for prediction as described in AHA guideline [14, 15].

## 4.2 Evaluation

This research also did comparison of three models: Naïve Bayes; Support Vector Machine (SVM); and deep learning. Algorithms Naïve Bayes and SVM are widely used in prediction. In Naïve Bayes, we used the classifier for discrete predictors that assumed independence of the predictor variables, and Gaussian distribution for metric prediction. The attributes with missing values and the corresponding table entries were omitted for prediction. In order to test, we calculated the values of the A-priori probabilities of 0.4784 (Non-Stroke) and 0.5215 (Stroke). Next, SVM method was used for sampling method in 10-fold cross validation for prediction, and SVM-Kernel was used for linear analysis. The cost of constraints or 'C'-constant of the regularization term is 0.1, and also their best performance is 0.465. Therefore, the prediction values of Naïve Bayes and SVM of 0 (Non-stroke) or 1 (Stroke). This result indicates that the patients are suffering from stroke or not.

In terms of deep learning, we used deep Neural Network model by using a feed-forward multi-layer artificial neural network. The computation shows that Mean Square Error (MSE) is 0.2596. This value indicates the confidence has best performance for prediction of stroke. The Mean Value and Standard Deviation are different from each technique, so the percentage of predict stroke with three models are showed in table 1. The result shows that during training procedure, the mean values for prediction in deep learning is higher than two techniques. For the result prediction of deep learning technique, it shows in percentage that different from Naïve Bayes and SVM which shows a chance of stroke. The number of edges with deep learning are plotted (see fig 3). An edge of weight value in each layer is the predictor with risk factors. So, a prediction of stroke implied some concepts of a medical domain that preferred for good performance and explanation.

**Table 1.** Comparison of three techniques used for prediction of Stroke.

Techniques	Mean Value	Standard Deviation
Naïve Bayes	49%	0.038
Support Vector Machine	47.78%	0.1106
Deep learning	36.73%	0.084

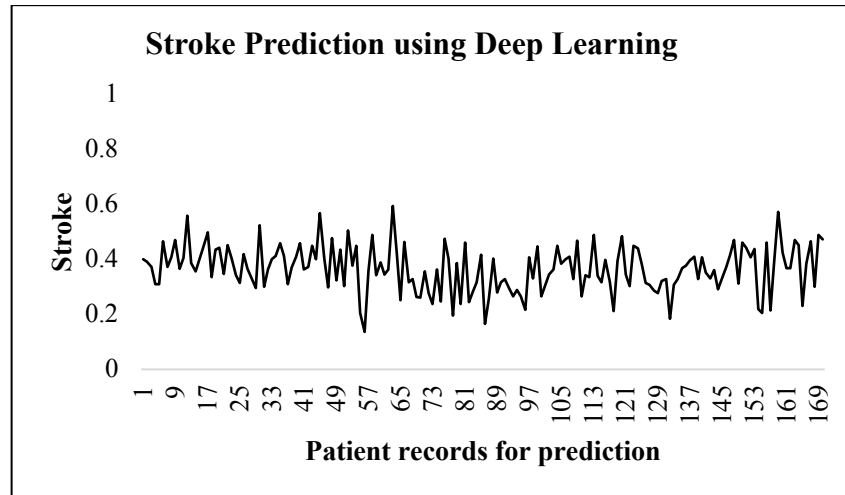


Fig. 3. Stroke prediction using deep learning.

## 5 Conclusion and Future work

In this paper, we compared three techniques: the deep learning technique; Naïve Bayes; and Support Vector Machine for stroke prediction. All techniques used data from heart patients for risk factors identification and disease prediction. The results are valid for warning to patients and significantly favorable to decision-making by a medical practitioner. We compared the use of deep learning with deep neural network, Naïve Bayes for discrete predictor, and linear performance in SVM. The result of Naïve Bayes and SVM show that patients are suffering from stroke or not deep learning technique shows in percentage of a chance of stroke. This confirmed that deep learning technique is most suitable for generating the heart dataset for predictive analysis in stroke.

Deep learning is widely used in prediction of diseases, especially in the prognosis and data analyses in healthcare sector. It employs the use of data obtained from patient health records and a comparison between previous cases, observation, or inspection. Stroke has complex risk factors. Therefore, the concepts or decision-making techniques cannot be directly extracted from a source that requires the involvement of a number of human experts.

However, a high volume of medical data, heterogeneity, and complexity have become the biggest challenges in diseases prediction. Algorithms with very high level of accuracy are, therefore, vital for medical diagnosis. The development of algorithms, nevertheless, still remains obscure despite its importance and necessity for healthcare. Good performance comes along with specific favourable circumstances, for instance, when well designed and formulated inputs are guaranteed. Consequently, the deep learning allows the disclosure of some unknown or unexpressed knowledge during prediction procedure, which is beneficial for decision-making in medical practice and provide useful suggestions and warnings to patient about unpredictable stroke. In

future, we will use more risk factors to predict in deep learning algorithm and compare to other techniques.

## References

1. Kononenko, I.: Inductive and Bayesian learning in medical diagnosis. *Applied Artificial Intelligence an International Journal* 7, 317-337 (1993)
2. Amin, S.U., Agarwal, K., Beg, R.: Genetic neural network based data mining in prediction of heart disease using risk factors. In: 2013 IEEE Conference on Information & Communication Technologies (ICT), pp. 1227-1231. (2013)
3. Leira, E.C., Ku-Chou, C., Davis, P.H., Clarke, W.R., Woolson, R.F., Hansen, M.D., Adams, H.P., Jr.: Can We Predict Early Recurrence in Acute Stroke? *Cerebrovascular Diseases* 18, 139-144 (2004)
4. Ohkubo, T., Asayama, K., Kikuya, M., Metoki, H., Hoshi, H., Hashimoto, J., Totsune, K., Satoh, H., Imai, Y.: How many times should blood pressure be measured at home for better prediction of stroke risk? Ten-year follow-up results from the Ohasama study. *Journal of Hypertension* 22, 1099-1104 (2004)
5. Khosla, A., Cao, Y., Lin, C.C.-Y., Chiu, H.-K., Hu, J., Lee, H.: An integrated machine learning approach to stroke prediction. *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 183-192. ACM, Washington, DC, USA (2010)
6. Letham, B., Rudin, C., McCormick, T.H., Madigan, D.: Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model. 1350-1371 (2015)
7. Sudha, A., Gayathri, P., Jaisankar, N.: Effective analysis and predictive model of stroke disease using classification methods. *International Journal of Computer Applications* 43, 26-31 (2012)
8. Liang, Z., Zhang, G., Huang, J.X., Hu, Q.V.: Deep learning for healthcare decision making with EMRs. In: 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 556-559. (2014)
9. Stier, N., Vincent, N., Liebeskind, D., Scalzo, F.: Deep learning of tissue fate features in acute ischemic stroke. In: 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1316-1321. (2015)
10. Nie, L., Wang, M., Zhang, L., Yan, S., Zhang, B., Chua, T.S.: Disease Inference from Health-Related Questions via Sparse Deep Learning. *IEEE Transactions on Knowledge and Data Engineering* 27, 2107-2119 (2015)
11. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* 521, 436-444 (2015)
12. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., Wang, Y.: Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology* (2017)
13. Russell, S.J., Norvig, P.: *Artificial intelligence: a modern approach* (International Edition). Pearson (2010)
14. Goldstein, L.B., Adams, R., Alberts, M.J., Appel, L.J., Brass, L.M., Bushnell, C.D., Culebras, A., DeGraba, T.J., Gorelick, P.B., Guyton, J.R., Hart, R.G., Howard, G., Kelly-Hayes, M., Nixon, J.V., Sacco, R.L.: Primary Prevention of



Ischemic Stroke: A Guideline From the American Heart Association/American Stroke Association Stroke Council: Cosponsored by the Atherosclerotic Peripheral Vascular Disease Interdisciplinary Working Group; Cardiovascular Nursing Council; Clinical Cardiology Council; Nutrition, Physical Activity, and Metabolism Council; and the Quality of Care and Outcomes Research Interdisciplinary Working Group: The American Academy of Neurology affirms the value of this guideline. *Stroke* 37, 1583-1633 (2006)

15. Goldstein, L.B., Adams, R., Becker, K., Furberg, C.D., Gorelick, P.B., Hademenos, G., Hill, M., Howard, G., Howard, V.J., Jacobs, B., Levine, S.R., Mosca, L., Sacco, R.L., Sherman, D.G., Wolf, P.A., del Zoppo, G.J., Members: Primary Prevention of Ischemic Stroke. *Circulation* 103, 163-182 (2001)