Long Short-Term Memory Recurrent Neural Network for Stroke Prediction

Pattanapong Chantamit-o-pas[1,2(✉)] and Madhu Goyal[1]

[1] Centre for Artificial Intelligence,
Faculty of Engineering and Information Technology,
University of Technology Sydney,
PO Box 123 Broadway, Ultimo, NSW 2007, Australia
Pattanapong.Chantamit-o-pas@student.uts.edu.au,
Madhu.Goyal-2@uts.edu.au
[2] Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang,
Ladkrabang, Bangkok 10520, Thailand
Pattanapong@it.kmitl.ac.th

Abstract. Electronic Healthcare Records (EHRs) describe the details about a patient's physical and mental health, diagnosis, lab results, treatments or patient care plan and so forth. Currently, the International Classification of Diseases, 10th Revision or ICD-10 code is used for representing each patient record. The huge amount of information in these records provides insights about the diagnosis and prediction of various diseases. Various data mining techniques are used for the analysis of data deriving from these patient records. Recurrent Neural Network (RNN) is a powerful and widely used technique in machine learning and bioinformatics. This research aims at the investigation of RNN with Long Short-Term Memory (LSTM) hidden units. The empirical research is intended to evaluate the ability of LSTMs to recognize patterns in multi-label classification of cerebrovascular symptoms or stroke. First, we integrated ICD-10 code into health record, as well as other potential risk factors within EHRs into the pattern and model for prediction. Next, we modelled the effectiveness of LSTMs for prediction of stroke based on healthcare records. The results show several strong baselines that include accuracy, recall, and F1 measure score.

Keywords: Deep learning · Cerebrovascular disease · Predictive technique
LSTM-RNN

1 Introduction

Deep learning algorithm is a technique that focuses on how computers learn from data. It is the intersection of statistics, computer science, and mathematics - which generates the algorithm of building patterns and models from massive data sets, as well as is applicable to billons or trillions of data records [1, 2]. Deep learning technique employs learning from data together with multiple levels of abstraction deriving from computational models that are associated with multiple processing layers.

Basically, a cerebrovascular disease or stroke is the state of lacking blood supply in an area of the brain, and it happens once a vessel is blocked. This is known as an "*ischemic stroke*", where about three-quarters of vessels are blocked. Meanwhile, a "*hemorrhagic stroke*" refers to the state where a blood vessel bursts. It can also affect different parts of the human body, depending on which area of brain is affected. In most countries, stroke becomes the second or third common cause of death [3, 4]. The patients who survived usually have poor quality of life because of serious illness, long-term disability and become burden to their families and health care system. This strongly demands for the management focusing on prevention and early treatment of diseases by analysing different factors. According to the analysis, it was found that several health conditions and lifestyle factors become risk factors for stroke.

The predictive techniques of stroke vary from simple to more complex models. The risk factors of stroke are complex and applicable to find different convolutions of disease and uncertainty from direct and/or indirect sources. The analysis of stroke patients who were admitted in the TOAST study was done by using stepwise regression methods [5]. This research was conducted among 1,266 stroke patients selected from database, provided that those patients must have had suffered a transient ischemic attack (TIA) or recurrent stroke within 3 months after the first stroke. Additionally, 20 clinical variables were chosen for finding performance and evaluation.

Some researches show that the use of ICD-9 codes in combination with other health care data can accurately diagnose patients' health issues [6–8]. Most of Electronic Healthcare Records (EHRs) adopt the codes from International Classification of Diseases (ICD), 10th Revision i.e. ICD-10 and ICD-10-CM codes, and those codes become the standard codification in the Electronic Medical Record system (EMR) [9]. The use of International Classification of Diseases in various health care institutions provides the similar basic schemes that allow patient data to be used in a similar way.

The rest of this paper is organized as follows. Section 2 describes related work on deep learning. Section 3 ICD-10 complaint Electronic Healthcare Records. Section 4 describes prediction of stroke using EHRs and deep learning. Section 5 presents the evaluation model. Section 6 discusses the result of this research and the conclusion and future work are present in Sect. 7 of this paper.

2 Deep Learning

Deep Learning method is intended to discover complex structure in big data set by using the advanced mathematical 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.

Basically, deep net has various techniques to predict the result. It is recommended to use either Restricted Boltzmann Machine (RBM) or auto encoder for unsupervised learning and the extraction of pattern from a set of unlabelled data. Several options are usable if there are labelled data for supervised learning, and once it is required to build classifier, depending on specific application. A recurrent network or Recursive Neural Tensor Network (RNTN) can be applied to text processing task like a sentence analysis based on phrasing, name and recognition. Moreover, Deep Belief Network (DBN) or

convolutional networks are used for image recognition. The RNTN or convolutional networks are also used for object recognition. Finally, a recurrent network is used for the speed recognition as well. In general, both DBN and Multilayer perceptron – also known as Rectified Linear Units (ReLUs) – are good choices for classification. Also, a recurrent network is the best option for time series analysis.

Gulshan et al. [10] invented an algorithm for automated detection of diabetic retinopathy in Retinal fundus photographs (RDR). In this research, a deep convolutional neural network was employed to optimize image classification and was trained by using a retrospective development dataset of 128,175 images. The results show that an algorithm has high sensitivity and specificity for detecting referable diabetic retinopathy.

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 endovascular c1otretrieval intervention is investigated. For the first time, Stier et al. [11] 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 over a single-voxel-based regression model. The previous researches reveal that the kernel of the deep learning techniques can be applied to healthcare sector as a regulariser at the output layers or a part of model.

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 et al. [12] suggested the application of deep belief network for unsupervised feature extraction, and then conducting 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.

In healthcare fields, data in EHRs are quite significant for decision-making in treatment. In general, a realistic dataset contains useful records for clinical practice. It uncovered realistic environments for the analyses of diseases because it had included ambiguous and incomplete values that contribute to errors and are unsuitable for annalistic data and a very challenging analysis. Normally, it needs to be fulfilled before being used. Hammerla et al. [13] proposed an assessment system that managed practical usability constraints and applied deep learning technique to differentiating disease state in datasets that are naturalistic settings. In this research, a large dataset was collected from 34 participants who suffered Parkinson's Disease (PD).

In other fields, deep learning is used in stock price prediction by extracting structure event from news text. The prediction technique uses event-driven approach. First, the system extracts the events from text and demonstrates dense vectors. Then, it trains

using a novel neural tensor network. Second, both short-term and long-term influences of event on stock price moments are combined and processed by using deep convolutional neural network. In comparison with state-of-the-art baseline methods, the results show that this model can achieve approximately 6% improvements on S&P 500 index prediction and individual stock prediction, respectively. In addition, market simulation results show that the system is more capable of making profits than previously reported systems trained on S&P 500 stock historical data [14, 15].

3   ICD-10 Complaint Electronic Healthcare Records

There is significant growth in the amount of medical or patient data being generated in hospitals or clinics all over the world. In most cases, Electronic Health Records are used for storing most of this medical or patient data. In practice, International Classification of Diseases (ICD) are used for electronic health records and for classifying diseases and other health problems appearing in many types of health and vital records. Currently, ICD-10 codes are used by hospitals and health professionals, which are retrieved from Electronic Medical Records (EMR) system. The standard provides a very convenient platform for primary and secondary data analysis of these records for diagnosis and prediction of diseases, as well as for the improvement of medical and patient care. The 'core' three character code of classification of ICD-10 is the mandatory level of coding for international reporting. It also has four character sub-categories which are not mandatory for international reporting [9]. The Electronic Healthcare Records (EHR) of cerebrovascular disease patients contain various information, including demographic data, potential risk factors, and non-potential risk factors that are recorded in hospital database (See Fig. 1).
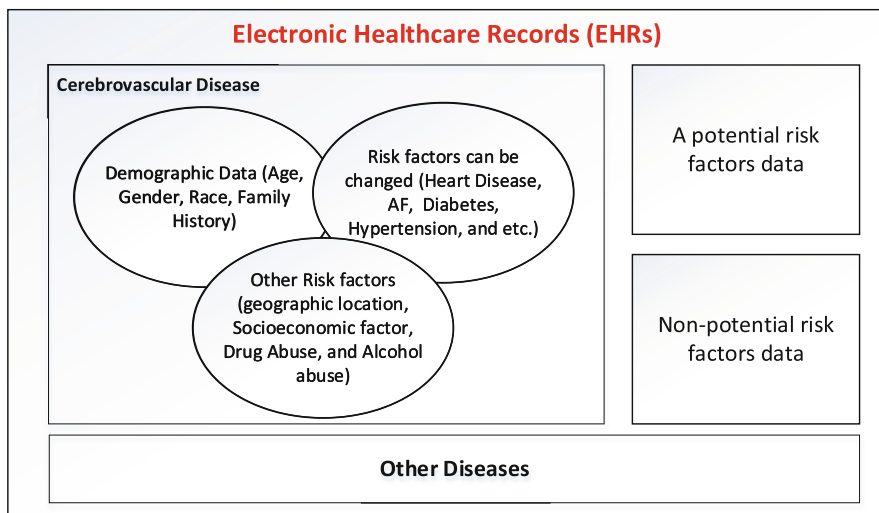

Fig. 1. Electronic healthcare records of stroke patients.

In detail, EHR record consists of gender; data of birth (DOB); clinic operation (CLINIC_OPD); date operation (DATEOPD); date diagnosis (DATEDX); clinic diagnosis (CLINIC_ODX); diagnosis code (DIAG); and diagnosis type (DXTYPE). Gender was identified and represented by either the code 1 (Man) or the code 2 (Woman) (see Table 1). DOB field record contains patient's birthdate and some records have null value or error in term of date. We eliminated the null value or error and converted the value into age in preparation process. CLINIC_OPD field indicates the clinic number of hospital where patient has treatment. DATEOPD field demonstrates the data of service. DATEDX field indicates the date of diagnosis. CLINIC_ODX field shares similar code with CLINIC_OPD field. Code of diagnoses were obtained from doctors or medical experts who used ICD-10 codes and inputted in DIAG field. DXTYPE field specifies types of disease. It consists of (1) primary disease, (2) co-morbidity disease, (3) complication, and (4) other diseases (see Table 1). Therefore, the demographic data, disease type, and other information were recorded once each patient visited. For prediction process, we integrated the multiple value dependencies into EHRs. Further details will be described in Sect. 4.

Table 1. Sample partial electronic healthcare records in hospital database.

Gender	DOB	CLINIC_OPD	DATEOPD	DATEDX	CLINIC_ODX	DIAG	DXTYPE
1	xx/xx/xxxx	120	xx/xx/xxxx	xx/xx/xxxx	120	C20	1
2	xx/xx/xxxx	20	xx/xx/xxxx	xx/xx/xxxx	20	D379	1
2	xx/xx/xxxx	10	xx/xx/xxxx	xx/xx/xxxx	10	C499	1
2	xx/xx/xxxx	120	xx/xx/xxxx	xx/xx/xxxx	120	K297	1
1	xx/xx/xxxx	10	xx/xx/xxxx	xx/xx/xxxx	10	C499	1
2	xx/xx/xxxx	10	xx/xx/xxxx	xx/xx/xxxx	10	I158	2
2	xx/xx/xxxx	120	xx/xx/xxxx	xx/xx/xxxx	120	I694	2

4 Prediction of Stroke Using EHRs and Deep Learning

In this paper, deep learning algorithm is applied on EHRs for prediction of stroke. Deep Learning (DL) is a process of training a neural network to perform given task. Overall the prediction of Stroke has two main steps i.e. selection of EHRs based on risk factors and prediction process. In the first step, first the null values and anomaly data were eliminated via JAVA programming. Then the ICD-10 codes were filtered by stroke's risk factors. The EHRs with ICD-10 codes are then filtered again to eliminate anomaly data such as negative values, null values etc. Then, EHRs files consisting of demographic data and group of symptom codes with risk factors are integrated. In the preparation phase, the EHRs were later converted to zero or one for defining diseases that the patients suffered (see Table 2).

In the second process, stroke is predicted by using Long Short-Term Memory - Recurrent Neural Network (LSTM-RNN), which is currently the most suitable approach. For prediction algorithm, the dataset was trained by means of feature selection and retrieval process, and LSTM-RNN prediction formula is applied. The input layer calculated the weight values based on ICD-10 codes and EHRs with risk

factors of stroke. The ICD-10 codes that represent stroke risk factors are selected by using AHA guideline [16–18]. This group was a knowledge-based reference that was used for computing the weight for embedding at hidden layer as input. The weight values and EHRs were integrated into LSTM-RNN layer. The output layer of prediction model represented the prediction value in a form of percentage risk (see Fig. 2).
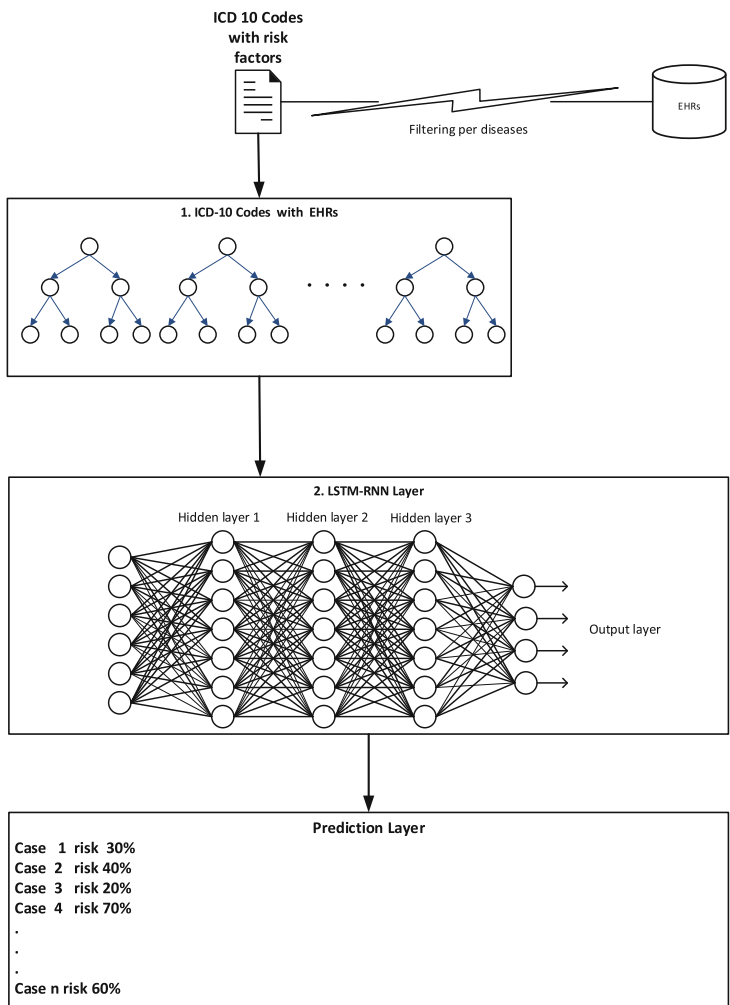

Fig. 2. Model of stroke prediction using EHRs and deep learning

4.1 The Selection of EHRs Based on Risk Factors

The ICD-10 code that had been applied in EHRs can be used for training probabilistic classifiers from the large data sets of EHRs. Specifically, we consider mulitlabel classification of stroke symptoms and risk factors for training and modelling by

selection based on AHA list of stroke factors [16–18]. Normally, ICD-10 code presented in main categories and sub-categories such as I65 (Occlusion and stenosis of vertebral artery) is the main category, and I65.1 (Occlusion and stenosis of basilar artery) is the sub-category. The code risk factor that has been chosen consists of 70 main-categories and about 200 sub-categories, all together are 227 factors.

In preparation process after cleaning the data, we reformatted the DOB filed by computed to age of patient. After filtering with ICD-10 codes, these codes will be shown in 1 or 0 to represent the existent of risk factor for each record. The former EHRs with mixed codes then was rearranged to a new EHRs dataset as show in Table 2. This new dataset is smaller in size and suitable for train and test in prediction process.

Table 2. EHRs records with stroke's risk factors.

Age	Gender	B980	E108	E119	……..	I64	Z721	Z920
74	1	1	0	0	……..	0	0	0
65	2	1	0	0	……..	0	0	0
61	1	1	0	1	……..	0	0	0
⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮
50	1	0	1	1	…….	0	0	0
62	2	0	1	0	…….	0	0	0
76	1	0	1	1	…….	1	0	0
83	2	0	0	1	…….	0	0	0

4.2   Deep Learning by Using Long Short -Term Memory Recurrent Neural Networks (LSTM-RNN)

In this section, we implemented the LSTM-RNN. An architecture contained computation units in each memory block in the recurrent hidden layer. The memory block contained memory cells with self-connections storing the temporal state of the network in addition to multiplicative unit that was called 'gate', which controlled the flow of information inputted to unit. An input gate and output gate were included in the original architecture. The input gate controlled the flow of information and activations into the cell that was computed by *sigmoid* and *tanh* function. The output gate controlled the output flow of cell that activation function was computed by using *sigmoid* and *tanh* function for the rest of the network.

The forget gate was added to the memory block. This gate prevented a weakness of LSTM models from processing continuous input streams that are not segments into subsequences. The internal state of cell of the forget gate scales run verification before adding an input to the cell through the self-recurrent connection of the cell; therefore, it would forget or reset the cell's memory [19]. This gate used *sigmoid* function for

computation. Furthermore, in the LSTM architecture peepholes connections (green line) form its internal cells were applied to all gates in the same cell for learning precise timing of the outputs [20] (see Fig. 3).
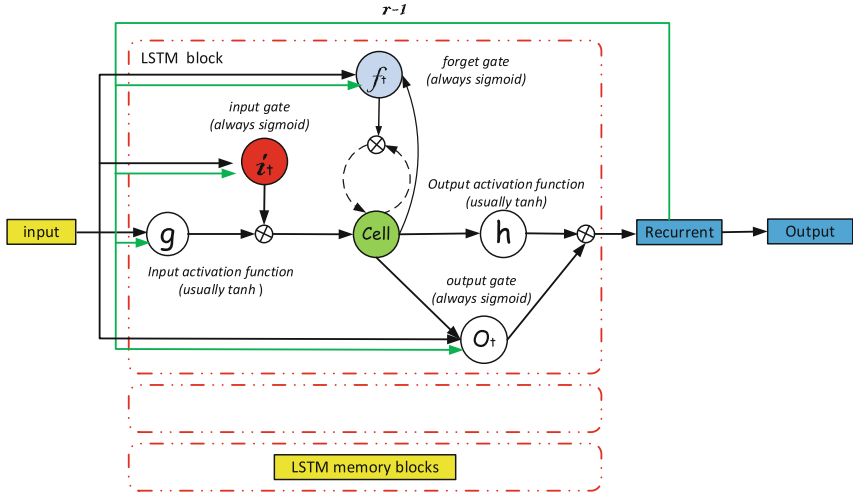

Fig. 3. LSTMP-RNN memory cell architecture and memory blocks [20–22].

5   Evaluation Model

5.1   Data Source

This research used aggregated files of Electronic Healthcare Records (EHRs) from Department of Medical Services, The Ministry of Public Health of Thailand between 2015 and 2016 (326,152 records). It consisted of demographic data, diseases codes (ICD-10 codes), Dates of diagnosis, clinic types, and types of diagnosis (see in Table 1). According to the source, EHRs data had multiple value dependencies that was cleaned anomaly data and filtered by ICD-10 codes for risk factors of stroke. Subsequently, we had a new EHRs dataset (See in Table 2). The new datasets actually had 96,127 records of the stroke patients and non-stroke patients who encountered potential risk factors.

5.2   Predictive Model

The algorithm deep learning relies Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) that is wildly used in prediction. In this research, it was applied to a large scale of an aggregated file from Electronic Healthcare Records. The algorithm model appears as follows:

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \tag{1}$$

$$a_i = \sigma(W_a x_i + W_a h_{i-1} + b_a) \tag{2}$$

$$g_i = \sigma\left(W_g x_i + W_g h_{i-1} + b_g\right) \tag{3}$$

$$\tilde{h}_i = tanh(W_h x_i + g_i \circ W_h h_{i-1} + b_h) \tag{4}$$

$$h_i = a_i \circ h_{i-1} + (1 - g_i) \circ \tilde{h}_i \tag{5}$$

Where the W terms are weight matrices value, W_h, W_b, and W_a, are diagonal weight values for next layer to connections. The b terms are bias vectors. The logistic sigmoid function is represented by σ. The input gate, forget gate, and output gate are represented by a, g, and n respectively. All of them are in the same size as the cell output activation vectors h_i, $\circ$ is the element product of the vector $\tilde{h}_i$ is the cell input and cell output activation function, generally and in this research network is *tanh*.

We initiated the prediction model for training stroke symptom and risk factors based on ICD-10 standard. The equations of the model appear as follows:

$$h_t = \tanh \frac{(W(I64 \sim Age) + W(I64 \sim Gender) +}{W(I64 \sim Stroke's\, risk\, factors))} \tag{6}$$

The machine learned from model and pattern. The group of codes was computed for finding the weight value in each node. The learning rate term is 0.01 and epoch in 10 and their network type is LSTM.

6 Result

We conducted a comparison between three models: Backpropagation; RNN; and LSTM- RNN. Algorithm backpropagation, RNN, and LSTM-RNN were applied in prediction. All techniques demonstrated the results of training 30%, 50%, and 80% respectively. For testing, 10% of the dataset is used. A learning rate is 0.1 and the number of iteration (10-epochs) were used for prediction. The variable for calculating was used for 227 risk factors for stroke.

The result shows that during training procedure, the accuracy value, precision value, recall value, and F1 scores for prediction in LSTM-RNN are higher than those obtained from the other two techniques. The result of RNN shown for all value lowest at 50% of sample size (0.3570; 0.3612; 0.6476; 0.5456).

In backpropagation, we used a feedforward multilayer artificial neural network. The computation shows that accuracy is at 0.8912, 0.8917, and 0.8914, F1 score as 0.3857, 0.3857, and 0.3860 respectively (shown in Table 3). This method show that all values are not differ in the three sample size for prediction. Only, the accuracy shown here is a little bit changed when the sample data slightly increased.

By using the same parameters as in previous techniques, the LSTM-RNN show the best results the best performance for prediction of stroke. The accuracy is 0.9279,

0.9493, and 0.9998 and F1 score are 0.9626, 0.9738, and 0.9999 respectively. The prediction for stroke will be shown in percentage which mean that change of getting stroke. In medical domain, the good performance is the preferred algorithm and LSTM-RNN is considered confidence to use with huge dataset.

Table 3. Metrics of stroke prediction

Train 30% and Test 10%

Techniques	Precision	Recall	F1	Accuracy
Backpropagation	1	1	0.3857	0.8912
RNN	0.9371	0.9375	0.9371	0.8825
LSTM-RNN	**0.0219**	**0.0216**	**0.9626**	**0.9279**

Train 50% and Test 10%

Techniques	Precision	Recall	F1	Accuracy
Backpropagation	1	1	0.3857	0.8917
RNN	0.3570	0.3612	0.6476	0.5456
LSTM-RNN	**0.9741**	**0.9738**	**0.9738**	**0.9493**

Train 80% and Test 10%

Techniques	Precision	Recall	F1	Accuracy
Backpropagation	1	1	0.3860	0.8914
RNN	0.9996	0.9996	0.9996	0.9992
LSTM-RNN	**0.9999**	**0.9999**	**0.9999**	**0.9998**

7   Conclusion and Future Work

This research aims at the use of deep learning technique and EHRs based on risk factors to predict for cerebrovascular disease by LSTM-RNN algorithm. The Electronic Healthcare Records (EHRs) provide the descriptive details about a patient's physical and mental health, diagnosis, lab results, treatments care plan and so forth. The data are difficult to mine effectively due to irregular sampling and missing data. Nowadays, the diagnoses of disease are represented by International Classification of Diseases, 10th Revision (ICD-10) code in each patient record. It enables researchers to train and develop a model to perform early diagnosis by predicting various risk factors.

The results of using LSTM-RNN show that accuracy rate, recall and F1 measure score are different from those of back propagation and RNN algorithm. Therefore, an accuracy rate depend on the size of sample. Unlike other techniques, the result is more reliable once there are large datasets for the prediction of stroke. This confirms that LSTM algorithm is most suitable for predictive analysis of any cerebrovascular disease or stroke.

EHRs using ICD-10 code have some issues and challenges in the data analyses of various diseases health problems by means of deep learning. The excellent analysis by different predicting techniques require the use of data obtained from patient health records and a comparison between previous cases, observation, or inspection. Stroke has complex risk factors, so 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. However, 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 can provide useful suggestions and warnings to patient about unpredictable stroke. In future, we will use more risk factors and lab results to predict using deep learning algorithm and implement to an e-stroke application.

References

1. Deo, R.C.: Machine learning in medicine. Circulation **132**, 1920 (2015)
2. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**, 436–444 (2015)
3. Langhorne, P., Bernhardt, J., Kwakkel, G.: Stroke rehabilitation. Lancet **377**, 1693–1702 (2011)
4. Khosla, A., Cao, Y., Lin, C.C.-Y., Chiu, H.-K., Hu, J., Lee, H.: An integrated machine learning approach to stroke prediction. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 183–192. ACM, Washington, D.C. (2010)
5. Leira, E.C., Ku-Chou, C., Davis, P.H., Clarke, W.R., Woolson, R.F., Hansen, M.D., Adams Jr., H.P.: Can we predict early recurrence in acute stroke? Cerebrovasc. Dis. **18**, 139–144 (2004)
6. Cooke, C.R., Joo, M.J., Anderson, S.M., Lee, T.A., Udris, E.M., Johnson, E., Au, D.H.: The validity of using ICD-9 codes and pharmacy records to identify patients with chronic obstructive pulmonary disease. BMC Health Serv. Res. **11**, 37 (2011)
7. Ried, L.D.P., Cameon, R.M.S., Jia, H.P., Findley, K., Hinojosa, M.S.P., Wang, X.P., Tueth, M.J.M.D.: Identifying veterans with acute strokes with high-specificity ICD-9 algorithm with VA automated records and Medicare claims data: a more complete picture. J. Rehabil. Res. Dev. **44**, 665–673 (2007)
8. Scheurer, D.B., Hicks, L.S., Cook, E.F., Schnipper, J.L.: Accuracy of ICD-9 coding for Clostridium difficile infections: a retrospective cohort. Epidemiol. Infect. **135**, 1010–1013 (2007)
9. WHO: ICD10: International Statistical Classification of Disease and Related Health Tenth Revision, vol. 2. World Health Organization, Geneva (2004)

10. Gulshan, V., Peng, L., Coram, M., et al.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA **316**, 2402–2410 (2016)

11. Stier, N., Vincent, N., Liebeskind, D., Scalzo, F.: Deep learning of tissue fate features in acute ischemic stroke. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2015, pp. 1316–1321 (2015)

12. Liang, Z., Zhang, G., Huang, J.X., Hu, Q.V.: Deep learning for healthcare decision making with EMRs. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 556–559 (2014)

13. Hammerla, N.Y., Fisher, J., Andras, P., Rochester, L., Walker, R., Plötz, T.: PD disease state assessment in naturalistic environments using deep learning. In: Twenty-Ninth AAAI Conference on Artificial Intelligence, pp. 1742–1748 (2015)

14. Ding, X., Zhang, Y., Liu, T., Duan, J.: Deep learning for event-driven stock prediction. In: Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015), pp. 2327–2333 (2015)

15. Ding, X., Zhang, Y., Liu, T., Duan, J.: Using structured events to predict stock price movement: an empirical investigation. In: EMNLP, pp. 1415–1425 (2014)

16. 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)

17. 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)

18. The American Heart Association: Comparison of 12 risk stratification schemes to predict stroke in patients with nonvalvular atrial fibrillation. Stroke **39**, 1901–1910 (2008)

19. Gers, F.A., Schmidhuber, J., Cummins, F.: Learning to forget: continual prediction with LSTM. Neural Comput. **12**, 2451–2471 (2000)

20. Gers, F.A., Schraudolph, N.N., Schmidhuber, J.: Learning precise timing with LSTM recurrent networks. J. Mach. Learn. Res. **3**, 115–143 (2002)

21. Sak, H., Senior, A., Beaufays, F.: Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: Fifteenth Annual Conference of the International Speech Communication Association, Singapore, pp. 338–342 (2014)

22. Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J.: LSTM: a search space odyssey. IEEE Trans. Neural Netw. Learn. Syst. **28**, 2222–2232 (2017)