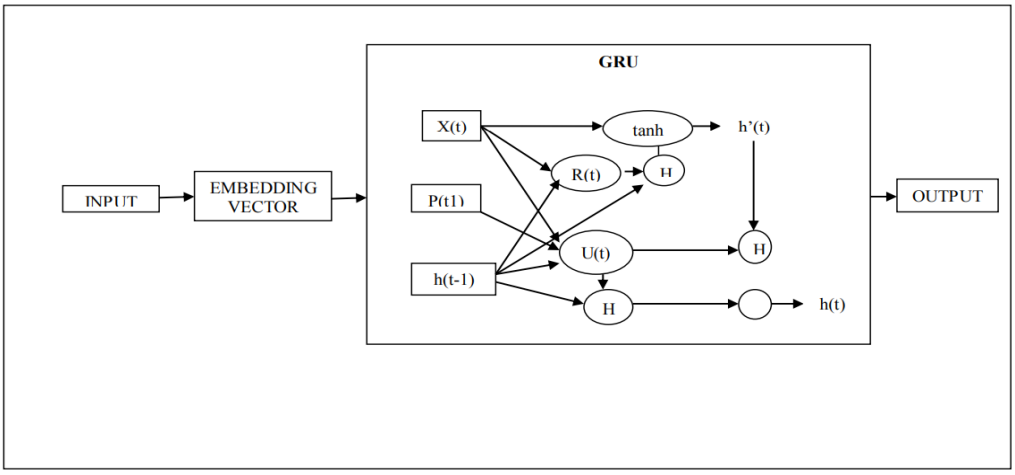
GRU Based Deep Learning Model for Prognosis Prediction of Disease Progression

Introduction

Recently different Deep Learning (DL) models have been emerging to predict the disease amelioration. Generally deep learning is the state of art for learning the multiple representation of neuron. Discriminant DL model includes different architecture and initial architecture is Recurrent Neural Network (RNN) which learn the data labeled in the form of sequence and it has long term dependency and vanishing gradient problem. Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) is an improved version of RNN which deal with problems in RNN. Deep care LSTM is one of the model in LSTM for predicting diabetes disease and compared the results with markovian and support vector machine model. In deep LSTM model some problems are reported like short term trajectories, less accuracy. To overcome the problem Deep Care GRU has been proposed to identify the diabetes disease amelioration.

Architecture of proposed work

The input is a sequence of admissions each admission contain the diagnoses code and intervention code explained about the disease and treatment histories of patients. The codes are embedded using vectors. The GRU architecture diagram for proposed work is illustrated in the following figure.



This approach is to embed admission into vectors. Let D be the set of diagnoses code indexed from 1 to D and I be the set of intervention code indexed from 1 to I. Each admission (t) contains a D: d1,d2……dn, ranges from 1 to D and I : p1,p2,……..pn ranges from 1 to P.

Update Gate

Update gate takes the previous illness history (h (t-1)) and current illness state (x(t)) and previous intervention histories p(t-1).

$$U(t) = \sigma(h(t-1) + x(t) + p(t-1)) \quad (1)$$

Reset gate

Reset gate takes h (t-1) and x (t) as the input and perform sigmoid operation, and it returns R (t).

$$R(t)\ \sigma(h(t-1) + x(t)) \quad (2)$$

Hadamard product is performed for same dimensional vector R (t) and h (t-1), and this result a (t) which determines what to be removed from the previous illness state. And e (t) is added with current input x (t). And now e(t) undergoes tanh activation and return h'(t)

$$h'(t) = \tanh(x(t) + (R(t) \odot h(t-1))) \quad (3)$$

Apply Hadamard function to 1-U (t) and h' (t) and also for U (t) and h (t-1) and sum the results to get the current illness state

$$h(t) = (1- U(t) \odot h'(t) + U(t) \odot h(t-1)) \quad (4)$$

Using this deep care GRU prognosis of diabetes disease have been predict and it also gives better accuracy compared to deep care LSTM model.

Experimental setup

Anaconda Navigator is the platform for different conda packages it is the Graphical User Interface (GUI) with many machine learning libraries and applications like jupyter notebook, spyder etc. Theano and Tensorflow are python based deep learning libraries whereas, theano is GPU based library and tensorflow is both GPU and CPU based library. It acts as a backend for Keras library. Keras is one of the deep learning library which supports for implementing prepackaged complicated architectures like RNN, LSTM and GRU etc. Tensorflow does not have many prepackaged architecture, it supports to design new architecture but keras support new datasets for known architecture. In this work keras and tensorflow libraries in jupyter notebook application is used. And also keras and tensorflow contains many prepackaged deep learning parameters like activation function, loss function. By using these parameters the architecture gives the better results. The parameters are explained as follows

Activation Function

Activation function is used to convert an input signal to the output signal. And this output signal is act as the input signal for the next layer. It uses weight function along with current state input and previous state output and process this parameter. The activation function is broadly classified into two types :

- Linear activation function
- Nonlinear activation function

Linear activation function defines that the output will not be inbounded in any specific range. Nonlinear activation function does not have any specific impound or range.

Sigmoid Function: Sigmoid is mainly used when the output is in the form of probability. That is the output is either 0 or 1. The sigmoid curve is in the form of S shape. The sigmoid function can be calculated using the below formula..

$$f(x) = 1 / 1 + \exp(-x) \quad (5)$$

Tanh function: Tanh is similar to sigmoid activation function, it ranges from (-1, 1).compare to sigmoid function tanh can have more gradient strength. But vanishing gradient problem still exist in both tanh and sigmoid function. The tanh activation function is illustrated in equation

$$f(x) = 1\text{-}\exp(-2x) / 1 + \exp(-2x) \quad (6)$$

Relu activation function: Relu is simple and efficient function. And Relu is more better than tanh activation function also it avoids vanishing gradient problem. The relu function is formulated in equation:

$$A(x) = \max(0, x) \quad (7)$$

Loss Function: Loss function is used to calculate the inconsistency between predicted value and actual value. The robustness of the model increases if loss function decreases.

Binary cross entropy: Binary cross entropy is a special case of categorical cross entropy. It uses only two classes for classification. The experimental determination of loss functions in using loss function.

Hidden Layers: Hidden layer is present between the input layer and output layer. The output of one hidden layer is the input of next layer. The hidden layer is a set of neurons which is used to process the input by adding some weight function.

Optimization Function: Optimization function is used to train the neural network. Optimization function helps to minimize the error.

Dataset Description

Diabetes dataset is used and 60% data is for training the data remaining is used for testing the model. The diabetes data for Deep GRU model gives the accuracy 70% and deep care LSTM gives 60% accuracy.

Conclusion

Deep learning can be used in many deep learning applications. Inn medical field the disease prognosis is predicted using deep learning. Here prognosis prediction for diabetes disease prediction was studied using Deep care GRU model. And the future progress of work is based on the Gated recurrent unit with its variants.