Survey on Text Error Detection using Deep Learning

Neethu S Kumar[1], Supriya L P[2]

[1]MTech, Dept. of Computer Science & Engineering, Sree Buddha College of Engineering, Pathanamthitta, Kerala
[2]Assistant Professor, Dept. of computer science & Engineering, Sree Buddha College of Engineering, Pathanamthitta, Kerala

---***---

Abstract – *Text Error detection is one of the important task of detecting errors in sentences. The existing methods to grammatical error correction involve rule-based and classifier approaches which are detecting only some type of errors in a sentence. A sentence may contain different types of errors and to detect the error is a difficult task. This paper describes a survey on text error detection using deep learning. This survey provides a brief introduction to the field and a quick of deep learning architectures and methods and existing text detection methods.*

Key Words: (Size 10 & Bold) Key word1, Key word2, Key word3, etc (Minimum 5 to 8 key words)…

1. INTRODUCTION

Error detection of text, which has wide application value, is an important research areaof natural language processing. Natural language processing (NLP) is a branch of machine learning that helps computers understand, interpret and manipulate human language. Natural language processing includes many different methods for interpreting human language. The basic tasks of natural language processing are tokenization and parsing, lemmatization or stemming, part-of-speech tagging, language detection and identification of semantic relationships. The field of natural language processing encompasses avariety of topics which involve the understanding of human languages. There are numerous complex deep learning based algorithms have been proposed to solve difficult NLP tasks. The major deep learning related models and methods applied to natural language tasks are convolutional neural networks (CNNs), recurrent neural networks (RNNs), and recursive neural networks.

Deep learning is also called feature learning or representation learning. Deep Learning is a set of machine learning algorithms which attempt to learn multiple-layered models of inputs, commonly neural networks. The main reasons to go deep is that a nonlinear function. It can be more efficiently represented by deep architecture with fewer parameters. Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans. n deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers. The term "deep" in deep learning refers to the number of hidden layers in the neural network. Traditional neural networks only contain 1-3 hidden layers, while deep networks contain more than 2 hidden hidden layers. Deep learning models are trained by large sets of labeled data. The neural network architectures that learn features directly from the data without the need for feature extraction. Deep learning is form of machine learning. A machine learning workflow starts with features manually extracted from images. The features are used to create a model that categorizes. With a deep learning workflow, features are automatically extracted from images.

2. DEEP LEARNING NETWORKS

In this section, discussed several deep learning networks such as Recursive Neural Network (RvNN), Recurrent Neural Network (RNN), Convolutional Neural Network(CNN), and deep generative models.

2.1 Recursive Neural Network (RvNN)

Recursive Neural Network[1] is one of the network of deep learning. It Uses a tree-like structure and preferred for Natural Language Processing. RvNN can classify the outputs as well as make predictions in a hierarchical structure using compositional vectors. The approach was to take a recursive data structure of variable size and generate a fixed-width distributed representation. RvNN has been especially successful in Natural Language Processing. The Back propagation Through Structure (BTS) learning scheme was introduced to train the neural network model [2]. It follows an approach similar to the standard back propagation algorithm and is also support a tree-like structure. The network is trained by auto association to reproduce the pattern of the input layer at the output layer.

2.2 Recurrent Neural Network(RNN)

Recurrent Neural Network is a simple type of Deep Neural Network. Recurrent neural network[3] is apply for sequence of information and it is one of the important algorithm of deep learning. This property is needful for many important applications such that finding word in a sentence, word embedding, text classification sentiment analysis and question and answering. RNN is mainly used for processing sequential information. The term recurrent state that it

perform same task at each instances. The recurrent neural network mainly three layers input layer , output layer and hidden layer. The input of the one node is the output of the previous node and the output of the one node is the input of the next node. The input of the network is in the form of vector. So it convert input into binary representation or machine languages. One hot encoded function is used for encoding. So the input of the network is encoded data. RNN have current time steps and which is depend on previous steps. At that time it make an update. So it need to back all the way to make updation. At that time may be many errors occurs due to the back propagation. RNN does not maintain the previous node. So it very complicated task to solve this problem because gradient of the loss function. And also called he vanishing gradient problem. To overcome this problem Long Short Term Memory(LSTM).

LSTM[4] is one of the funny neural network model which is based on recurrent neural network. The LSTM have an extra piece information than recurrent neural network which is called memory cell. It contain a piece of information of each node stored at each block of memory. Each memory block contain input gate, output gate and forget gate. The input gate controls the input activation. The output gate controls the cell activation.

2.3 Convolutional Neural Network(CNN)

Convolutional Neural Network[5] is one of the most important Deep learning algorithm. It can be used in many applications such as Natural Language Processing[6], Speech Recognition[7] and Computer Vision[8]. CNN has mainly three advantages. They are parameter sharing, sparse interactions, and equivalent representations. In Convolutional neural network contains convolutional layers and followed by subsampling layers and in the last stage, fully connected layers are used. CNN are commonly used in image and video processing. CNNs are capable of learning features and it may be present in different regions of the input data.

3. CONCLUSIONS

Deep Learning is one of the important new and hot topic of machine learning. It defined as deeply connected neural networks. Deep learning is mainly used to performing nonlinear processing and also processing multiple level of data representations. Deep learning has difficulty in modeling multiple complex data. Deep learning needs datasets for training the machine and predicting the unseen data. The existing deep learning implementations are supervised algorithms,

while machine learning is gradually shifting to unsupervised and semi supervised learning to handle real data with human labels.

REFERENCES

[1] Richard Socher, Cliff C. Lin, Chris Manning, and Andrew Y. Ng. 2011. Parsing natural scenes and natural language with recursive neural networks. In International Conference on Machine Learning. Omni press, 129–136

[2] Christoph Goller and Andreas Kuchler. 1996. Learning task-dependent distributed representations by back propagation through structure. In IEEE International Conference on Neural Networks, Vol. 1. IEEE, 347–352.

[3] Kyunghyun Cho, Bart van Merrienboer, Çaglar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In The Conference on Empirical Methods in Natural Language Processing. 1724–1734.

[4] Xiangang Li and Xihong Wu. 2015. Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 4520–4524.

[5] Yann LeCun and Yoshua Bengio. 1995. Convolutional networks for images, speech, and time series. Handbook of Brain Theory and Neural Networks 3361, 10 (1995), 255–257.

[6] Zhiwei Zhao and Youzheng Wu. 2016. Attention-based convolutional neural networks for sentence classification.In The 17th Annual Conference of the International Speech Communication Association. ISCA, 705–709.

[7] George E. Dahl, Dong Yu, Li Deng, and Alex Acero. 2012. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. IEEE Transactions on Audio, Speech, and Language Processing 20, 1 (2012), 30–42.

[8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.). Curran Associates, 1097–1105.

BIOGRAPHIES

Neethu S Kumar received the Bachelor's Degree in Computer Science and Engineering from Sree Buddha college of Engineering, Kerala, India in 2017. She is currently pursing Master's Degree
in Computer Science and Engineering in Sree Buddha College of Engineering, Kerala, India

Prof. Supriya L P. has more than 12 years of experience in teaching, Research and industry. She completed her post-graduation in Computer Science from Madras University in 2003. She received her M.Phil. From the department of computer Science in 2007, Annamalai University specialized in image processing .She received her Master of Engineering (M.E) degree from School of Computing, Sathyabama University, Computer Science and Engineering in 2009. At present she is pursuing her PhD. She started her career as a faculty of Computer Science in 2004 at Chennai. She has got a number of publications in conferences and Journals national/international.