Review of Week 1

Xiaoke(Jimmy) Shen

September 22, 2017

1   Introduction

This paper review report is provided based on the requirement of the computer science research course provided by the Graduate Center of the City University of New York during the 2017 Fall Semester. The main objective of this course is helping the PHD candidates to prepare their second exam and identify their thesis topic as early as possible [1].

2   Deep Learning Theory

2.1   Approximation with Artificial Neural Networks [2]

In order to build the mathematical theory of the artificial neural networks, several papers are published in the 20 century. In this paper one main contribution is the universal approximation theorem with proof. The universal approximation theorem claims [2] that the standard multilayer feed-forward networks with a single hidden layer that contains finite number of hidden neurons, and with arbitrary activation function are universal approximators in $C(R^m)$. The universal approximation theorem is one of the important theoretical support for the artificial neural networks. However, at that time as the huge size labeled data sets are not available, the updated algorithms haven't been invented and also the limited computation power, these ideas can not be verified.

2.2 Approximation Capabilities of Multilayer Feedforward Networks [3]

The unique value of the Kurt Hornik (1991) [3] paper is it showed that it is not the specific choice of the activation function, but rather the multilayer feedforward architecture itself which gives neural networks the potential of being universal approximators. This is an important contribution which is the foundation for the current state of the art deep learning architecture such as VGG 16 [4] and resnet [5]

2.3 Learning representations by back-propagating errors [6]

The paper of 1986 of BP(Back Propagation) significantly contributed to the Artificial Neural Networks [6]. It experimentally demonstrated the usefulness of the internal representations of hidden layers. The Back Propagation algorithm is one of the most critical and fundamental algorithm used in the deep neural network. It gave an efficient way to update the weights of the Neural Network. However, as the limitation of the computation power and also some other unresolved problems at that time such as overfitting and big number of labeled data is not available, the importance of this algorithms is still not well known as it is shown recently.

References

[1] P. Ji, "Syllabus of the computer science research at cuny fall 2017," 2017.

[2] B. C. Csji, "Approximation with artificial neural networks," pp. 11–12, 2001.

[3] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. vol. 4, 1991.

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.

[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 10 1986.