Deep Learning for Knowledge Graph Completion

Nick, Kevin, Cesar Bartolo-Perez, Doug

University of California - Davis

November 30th, 2017

Abstract

Contents

1 Introduction

An example of citing something is [1]

2 Methods

Multilayer Perceptron Model (MLP)

The multiyaler perceptron (MLP) model was built using the the Stuttgart Neural Network Simulator implemented in language R (RSNNS)[**?**]. Where we have used standard backpropagation and backpropagation with momentum methods in our simulation models.

Different MLP architectures are built, varying number of hidden layers, nodes and iterations in order to improve the accuracy and precision of our model.

Weighted SSE by number of iterations of the training and tetsing sets in addition to ROC curves will provide the references in our model architecture selection.

3 Results

In order to prevent overfitting in our model, we have compared our model with 100 and 50 interations (Figure **??**). We can notice that under 100 iterations, the trainin error drops to a weigthed SSE of almost 500. However, our testing error, start increasing after 30 itearions aproximately. From this results, we can infer that we have reached an overfitting in our model and a smaller number of iterations are enougth. In that sense, we have run our model with 50 iterations, avoiding the increase in the SSE.

Figure 1: MSE at different iterations. The structure presents a single hidden layer with 128 nodes
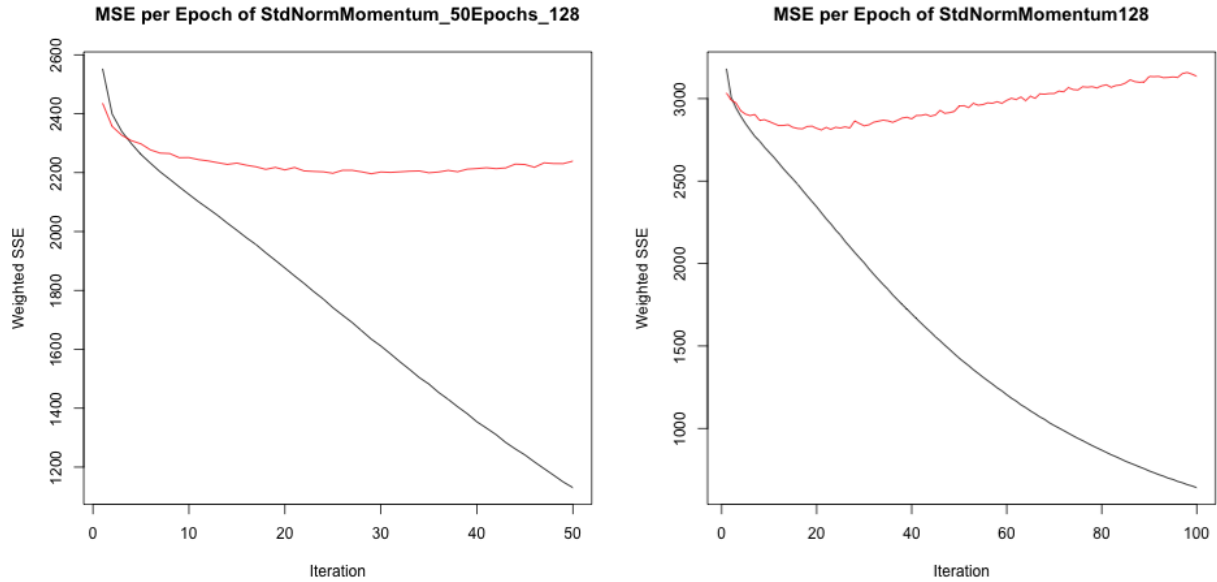
In a second running of the model, we have added a new hidden layer with different number of nodes (16,64 and 128). Before 50 iterations, adding nodes to the second hidden layer, reduce the error in our model. However, after 50 iteration, the higher number of nodes, increase rapidly the overfitting of the model.

Figure 2: MSE for A NN with two hidden layers with different nodes at second hidden layer.
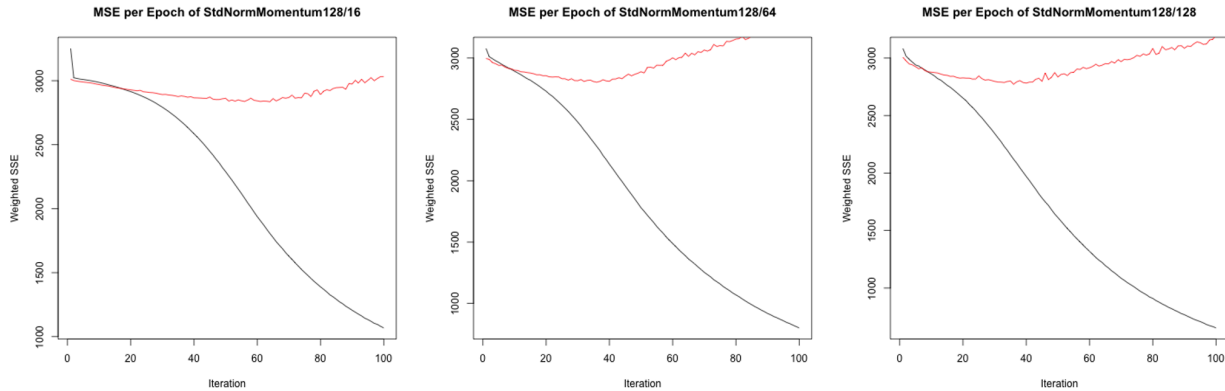

In a following set of tests, we compared different Neural Networks structures, increasing the number of hidden layers (hidden layers/nodes: 128, 128/16 and 128/128/16). Figure **??** shows that for the model with three hidden layers the training error is not decreasing considerably before 40 iteration. Suddently, the SSE in this NN structure drops abruptaly. The training error, shows that the error keep increasing at that range of iterations. This beahvior, can be attributed to an overfitting in our model.

Ths trend described above is similar for two hidden layers. For teh case of just one hidden layer with 128 nodes, the error does not drop abruptaly for the training set but we can see an increase in the error for after 30 iterations aproximately.

These results, make us suggest that only one hidden layer is enough for our MLP model, since there is not a big improvement in the reduction of error with more layers added and with the benefit of less computational resources are required. This decision is confirmed comparing their ROC curves for each NN structure.

Figure 3: MSE and ROC curves for different levels of hidden layers
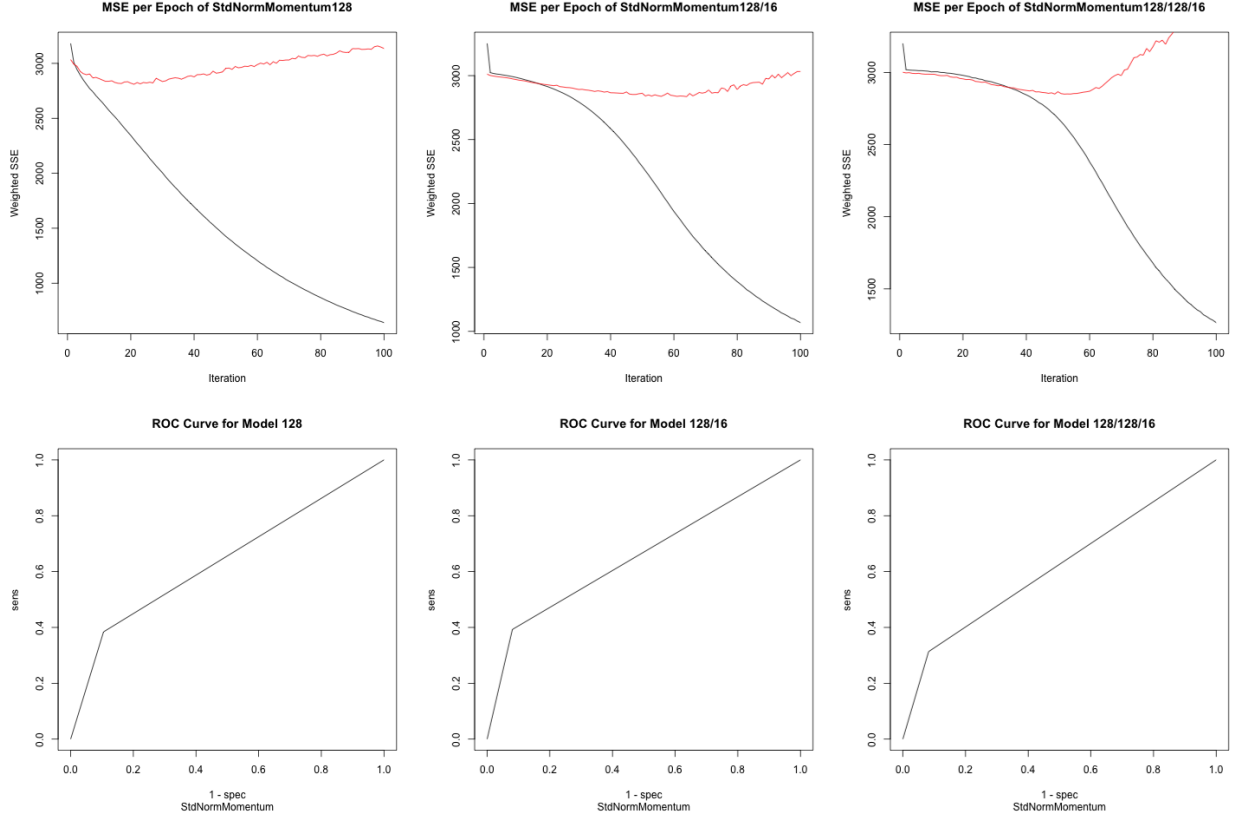
4   Discussion

References

[1] A. Kostic, D. Gevers, C. Pedamallu, M. Michaud, F. Duke, and A. Earl, "Genomic analysis identifies association of fusobacterium with colorectal carcinoma," Genome Res, vol. 22, pp. 292–298, 2012.