Assignment 5.1: Wine Quality Prediction Using Boltzmann Machines

Patricia Enrique

Department of Engineering, University of San Diego

AAI-511: Neural Networks and Learning

Prof. Esnaeilli

July 31, 2023

Introduction

Boltzmann machines are unsupervised deep learning models. Their architecture consists of hidden and visible nodes where each node is connected to every other node with an undirected connection. For this analysis, the Bernoulli Restricted Boltzmann machine model from scikit learn is used (Scikit Learn, 2023). The Restricted Boltzmann machine has only two layers, a visible layer and a hidden layer, where none of the visible nodes are connected to each other, and none of the hidden nodes are connected to each other. Due to the node connections being undirected, the weights are updated through contrastive divergence (Goodfellow et al., 2016).

Results

The results from experimenting with different parameters in the Restricted Boltzmann Machine are summarized in the tables that follow. For each model with parameter tuning, the loss at each epoch iteration is calculated. No model experiences a large change in loss ensuring there is no overfitting. From Table 1 it can be seen that the performance of the models remains consistent regardless of the number of hidden layers added. The wines that receive a good rating outperformed the wines with a poor rating.

Table 1

Exploration of hidden layers

Performance Metric	100		250		500	
	0	1	0	1	0	1
Accuracy	0.72		0.73		0.73	
Precision	0.67	0.74	0.66	0.75	0.66	0.75
Recall	0.47	0.86	0.52	0.85	0.52	0.85
F1	0.55	0.80	0.58	0.80	0.58	0.80

Table 2 summarizes the results from experimenting with different learning rates. A learning rate of 0.001 achieves the best accuracy while the precision remains similar for the three learning rates. Recall and F1 scores for the wines with a poor quality have the worst performance from the metrics with a learning rate of 0.01 and 0.1. Although the results are not optimal, a learning rate of 0.001 obtained the best results when comparing the poor wine quality performance with the good wine quality performance.

Table 2

Exploration of learning rates

Performance Metric	0.001		0.01		0.1	
	0	1	0	1	0	1
Accuracy	0.72		0.66		0.64	
Precision	0.66	0.75	0.75	0.66	0.88	0.63
Recall	0.51	0.85	0.13	0.97	0.02	1.00
F1	0.58	0.79	0.22	0.78	0.03	0.78

Table 3 summarizes the results from experimenting with different batch sizes. The model performance achieves the same results with a batch size of 10 and 50. When comparing the batch size of 1 and 10, the recall and F1 scores for the poor-quality wine are drastically better with a batch size of 10.

Table 3

Exploration of batch size

Performance Metric	1		10		50	
	0	1	0	1	0	1
Accuracy	0.66		0.72		0.72	
Precision	0.73	0.66	0.66	0.75	0.66	0.75
Recall	0.14	0.97	0.51	0.85	0.52	0.84
F1	0.24	0.78	0.58	0.79	0.58	0.79

The final model uses the trends found from the parameter exploration to optimize the performance. The model consists of 300 hidden layers, a learning rate of 0.001, and a batch size of 10.

Table 4

Final model performance metrics on test dataset

Performance Metric	Test Dataset	
	0	1
Accuracy	0.73	
Precision	0.64	0.77
Recall	0.55	0.83
F1	0.59	0.80

Analysis

From the three parameters chosen to evaluate, the learning rate is the one that has the most impact on the model's performance. The accuracy and precision metrics remained consistent and between 63% and 75% with an outlier of 88% regardless of the changes made to the model's parameters. The recall and F1 scores have the largest difference in performance between the good and bad quality wines.

The instances with a good quality rating consistently outperform the instances with a poor-quality rating, implying the possibility of an imbalanced dataset. A technique that can be used to mitigate this issue is oversampling and under-sampling the classes during the preprocessing stage.

The complete dataset is generated by combining a red wine and a white wine dataset. Since the goal is to determine the quality of the wine and not the type, the red and white labels were dropped. However, the addition of another feature classifying whether the wine is red or white can be beneficial in improving the model performance.

References

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. *MIT Press*.

Scikit Learn. (2023, July 30). Restricted Boltzmann Machine features for digit classification. https://scikit-learn.org/stable/auto_examples/neural_networks/plot_rbm_logistic_classification.html