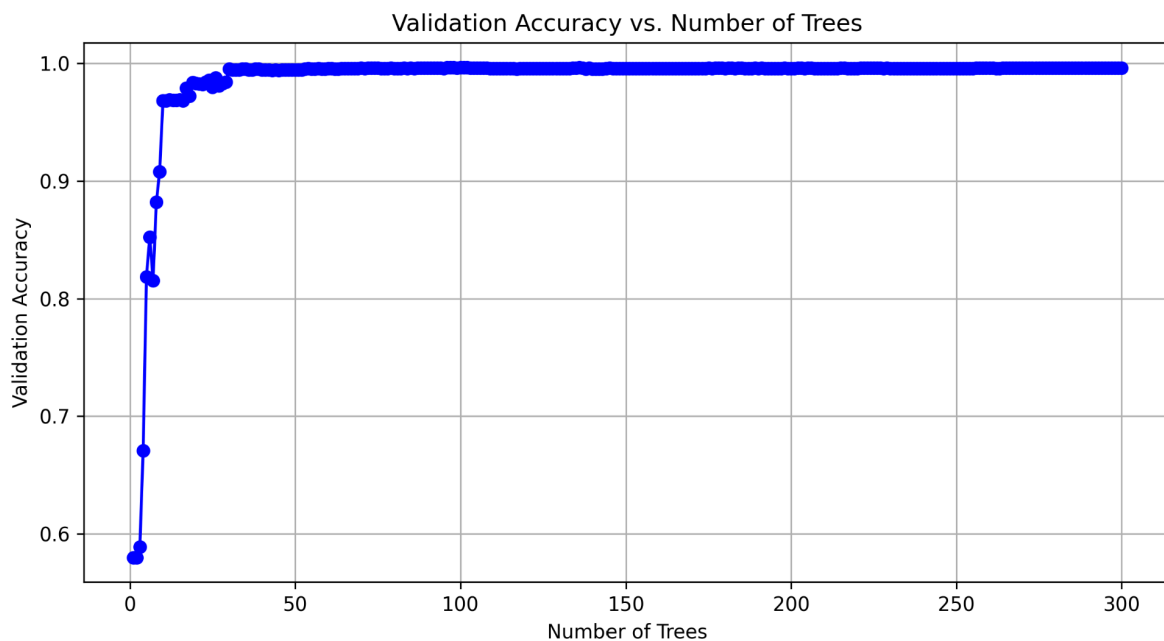


SML Assignment-4 Report

After basic data pre-processing as instructed in the assignment and dimensionality reduction ($n=5$), we proceed to construct the AdaBoost Classifier.

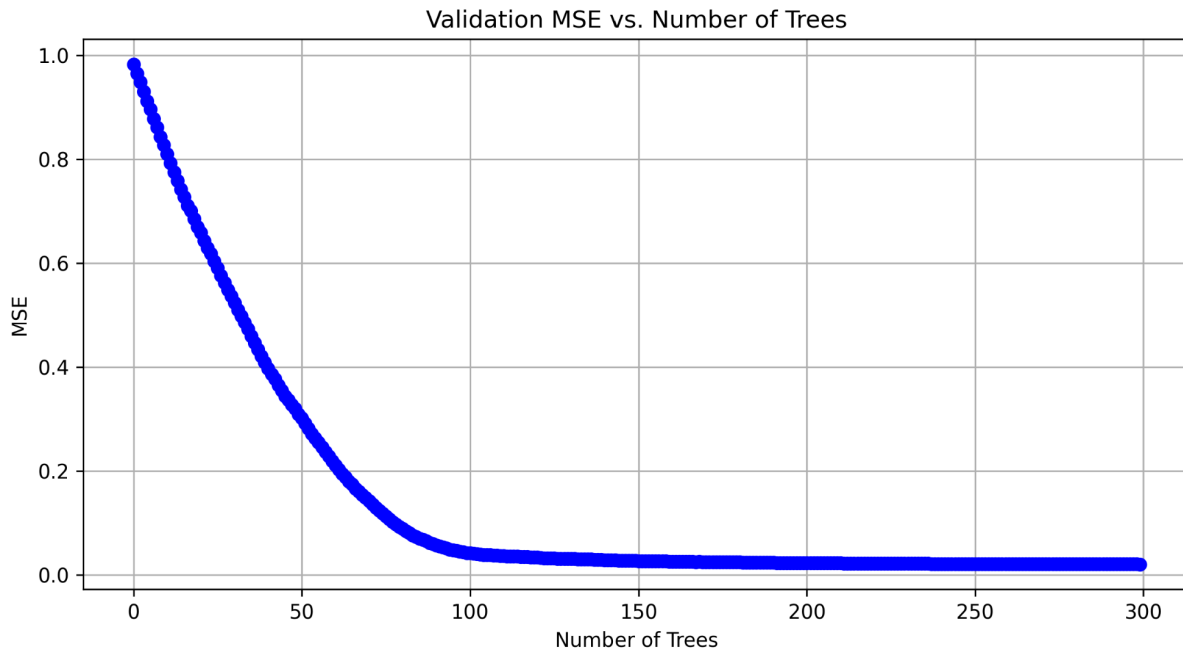
Each stump has been allowed to select features randomly. The stumps are evaluated based on the weighted classification error and their performance is measured by calculating the alpha values. The weights of the training samples are updated iteratively.

The model achieving the highest validation accuracy is considered as the best model and is used to calculate accuracy on the unseen test set.



```
Best tree count: 96, Validation Accuracy: 0.9965  
Test Set Accuracy: 0.9962174940898345
```

For Q2, each stump attempts to minimize the mean squared error (MSE) between the predicted and actual labels. The influence of each stump was controlled by a learning rate, ensuring that each stump makes a conservative contribution to the final model.



```
Best model at tree 300 with MSE 0.020556400000000002
Test MSE: 0.01232548463356974
```