ECEN 689 : Project 2

Samyuktha Sankaran, Harish Kumar, PRS Harinath

Abstract

Ensemble methods can take a large number of weak learners and aggregate them into a strong classifier. The outputs of the various weak learners are often aggregated either by voting in case of classification or by averaging in case of regression. In this project, we first examine whether strong learners can be aggregated in a similar manner to produce an ensemble classifier with higher performance than a single strong learner with equivalent complexity. We also demonstrate that training the set of individual learners is computationally less expensive than training a single model with equivalent performance.

I. PROBLEM DESCRIPTION

THE problem involves finding whether strong learners can be combined to produce an ensemble classifier that is ”better” than individual classifiers of far greater model complexity. We enforce a constraint that the combined complexity of all the individual learners in the ensemble is similar to the complexity of the individual classifier that we are comparing against.

II. TASK DIVISION

- Literature survey(Samyuktha)
- Data pre-processing (Samyuktha and Harinath)
- We use three different approaches for aggregating the outputs of the individual neural networks(1 approach/person).
- Comparing these results to an equally complex network(1 each).
- Theoretical analysis(using PAC learnability theory) and bounds to demonstrate ensemble of neural networks has better results than a single unit. (Harish Kumar)
- Attempt the same on multiple datasets.

III. TIME LINE

- Week 1 (November 25 - December 1): Complete Literature survey and finalize of the data sets. Preprocessing of datasets.
- Week 2 (December 2 - December 8): Building ensemble models and testing on the datasets.
- Week 3 (December 9 - December 12): Finalizing results and project report.