Spaceship Titanic: Applying Stacking Methods to a Kaggle Competition

Jordan Boswell

[jboswell30@gatech.edu](mailto:jboswell30@gatech.edu)

Abstract

Stacking is an effective ensemble method, yet there is a lack of literature exploring its best practices and implementation details. In this paper, we explore the effectiveness of various stacking and non-stacking models when applied to the Spaceship Titanic classification competition of Kaggle. The results show that stacking is generally better than single models, but the effectiveness of stacking depends on the training method used.

Introduction

Stacking, first proposed by Wolpert in 1992 [1], is an ensemble method that combines multiple strong base models into a meta model. The main idea is to use the predictions of the base models as input predictors to the meta model. The meta model, recommended to be from the family of regularized linear models [2], works best when its predictors are uncorrelated

Data

Methodology

Results

Conclusions

Lessons Learned

Figures

References