

Introduction

- Typical Random Forests and other similar methods split along coordinate axis
- Further developments created algorithms that split on linear combinations of coordinates (oblique), which allow better learning at the expense of computational efficiency
- SPORF uses sparse random projections for splits

Background & Related Work

- RF builds T decision trees on a random subset of the features, then the random forest makes a prediction based off of the most popular decision tree prediction
- Oblique extensions try to relax restriction for splits to be along coordinate axes of feature space
- Very sparse random projections can be used
- Gradient boosted trees use gradient descent iteratively
- Desirable properties for tree ensembles
 - Random search of splits
 - Flexible sparsity
 - Ease of tuning
 - Data insight
 - Expediency and scalability

Methods

- SPORF samples in a less restrictive way than RF