Abstract

Classification of high dimensional data is always a great challenge because of "small n, large p" or "the curse of dimensionality" like problems. To classify high dimensional data simply as it is is not feasible because none of existing classifiers can handle even tens of thousands in magnitude of dimentions. So we need to use some sophisticated methods in order to work with high dimentional data. First, we need method to drastically reduce dimentionality. Second, there is a natural need to identify stable subset of features that would be user for classification. Third, we need a technique that would be fast and scalable. And the ultimate goal is to have a good classification performance.

1 Introduction

1.1 Subtitle

A feature subset selection is a task of choosing a small subset of features that ideally is necessary and sufficient to describe the target concept.

1.2 Another subtitle

More plain text.

2 Experiments and Results

2.1 Data Sets

We user colon data set.

2.2 Experimental Setup

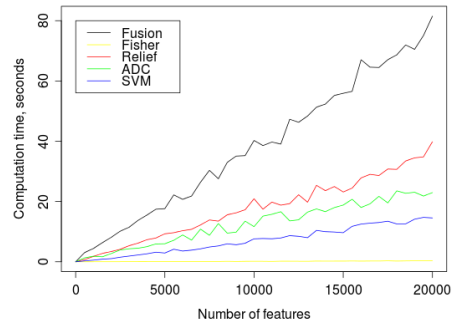For classification we used linear SVM

2.3 Results

moo.

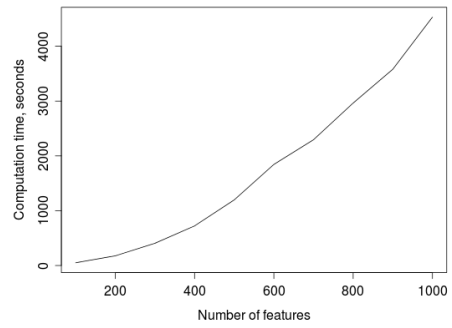Figure 1: Feature selection speed performance of various methods.

Figure 2: Conensus groups stable feature performance.