UNIVERSITAT POLITÈCNICA DE CATALUNYA

FACULTAT D'INFORMÀTICA DE BARCELONA

Extension Report: Dynamic Step

Analysis of the SVM-RFE algorithm for feature selection

Author:
Robert PLANAS

Director:
Luis A. BELANCHE

Bachelor Degree in Informatics Engineering
Specialization: Computing

March 28, 2021

Contents

Chapter 1

Dynamic Step

This modification is based on the constant step variant of the SVM-RFE algorithm, however, instead of using some constant number as the step in each iteration we calculate that number dynamically. The most straightforward way to do this is by using a percentage. Another possibility is to let the model score influence this percentage.

1.1 Description and reasoning

1.1.1 Method 1: Percentage

The percentage is a hyperparameter. It is used within every iteration to select a number of the first ranked features. A constant step has already been used in practice, but it is expected that this method will be significantly faster without effecting the accuracy performance, or even improving it.

The reasoning behind this is that when you have many features for which you only want to select a very small subset, the required amount of iterations will be of lineal complexity on the amount of features, however, by using a percentage, we get a logarithmic complexity.

We assume that, the bigger the step in each iteration, the worse the performance of the final selection. However, since we're selecting the worst variables first, selecting more of them at once shouldn't affect performance because they would have been selected anyway with height probability in the following iterations. The fewer the iterations remaining, the riskier it becomes selecting multiple variables at once, and thus a smaller step is beneficial.

1.2 Pseudocode formalization

Definitions:

- $X_0 = [\vec{x_0}, \vec{x_1}, \ldots, \vec{x_k}]^T$ list of observations.
- $\vec{y} = [y_1, y_2, \ldots, y_k]^T$ list of tags.
- F list of features.
- S list of selected features.
- S_f final selection, with $m = |S_f|$.
- A list of non-selected features.
- p the *percentage* hyperparameter.

Algorithm 1: SVM-RFE

Output: $\vec{r}$
Data: $X_0, \vec{y}$
$\vec{s} = [1, 2, \ldots, n]$ `// subset of surviving features`
$\vec{r} = []$ `// feature ranked list`
while $|\vec{s}| > 0$ **do**
    `/* Restrict training examples to good feature indices */`
    $X = X_0(:, \vec{s})$
    `/* Train the classifier */`
    $\vec{\alpha} =$ `SVM-train`(X, y)
    `/* Compute the weight vector of dimension length` $|\vec{s}|$ `*/`
    $\vec{w} = \sum_k \vec{\alpha}_k \vec{y}_k \vec{x}_k$
    `/* Compute the ranking criteria */`
    $\vec{c} = [(w_i)^2 \text{ for all } i]$
    `/* Find the feature with the smallest ranking criterion */`
    $f =$ `argmin`$(\vec{c})$
    `/* Update the feature ranking list */`
    $\vec{r} = [\vec{s}(f), ...\vec{r}]$
    `/* Eliminate the feature with smallest ranking criterion */`
    $\vec{s} = [...\vec{s}(1 : f-1), ...\vec{s}(f+1 : |\vec{s}|)]$
end

1.3 Expectations