



## CPS Paper

### Probabilistic Vector Machines

**Author:** Prof. Antonio Pedro Duarte Silva

**Submission ID:** 697

**Reference Number:** 697

#### Presentation File

[abstracts/ottawa-2023\\_03dae83d2fa5d5e62ad48301d0d70258.pdf](https://www.isi2023.org/abstracts/ottawa-2023_03dae83d2fa5d5e62ad48301d0d70258.pdf)

#### Files/Uploads

[ISI2023](#)

#### Brief Description

In supervised classification problems, Support Vector Machines (SVMs) are known to be among the most accurate class predictors.

However, standard SVMs fail to complement these predictions with reliable estimates of class membership probabilities.

Here, it will be presented a novel algorithm to estimate class probabilities from sequences of weighted SVMs.

Numerical experiments show that this algorithm, scales better than existing alternatives, is more accurate than competing machine learning approaches, and more robust than model based statistical methodologies

#### Abstract

Over the last few decades, kernel based methods have added an important set of tools to any professional statistician providing competitive, model free, alternatives to traditional statistical methodologies. In particular, in two-class supervised classification problems, kernel based Support Vector Machines (SVMs) are known to be among the most accurate predictors of unknown class memberships, and sequences of weighted SVMs can be used to accurately estimate the corresponding class probabilities. However, existing global all-in-one extensions of this approach to multiclass problems can be computationally too demanding, and do not scale well when the number of different classes grows. In this work, we will present an improved method to build reliable k-class probability estimates from global weighted SVMs with good scaling properties as k increases. Numerical experiments show that class probability estimation based on weighted SVMs is often more accurate than competing distribution free machine learning approaches, and more reliable than model based statistical methodologies when their assumptions can not be guaranteed. A public domain R package implementing the methods proposed in this paper is under preparation.