DOI 10.1007/s10994-013-5425-9

EDITORIAL

Enhancing Support Vector Machines with Fuzzy M-Estimator

Inspired Approaches for Robust Classification

# Min Gyeong Kim 1,† · Jia Yoo 1,† · Jin Hee Yoon1,\*

Received: 17 October 2013 / Accepted: 15 November 2013 / Published online: 28 November 2013

© The Author(s) 2013

**Abstract**

Support Vector Machines (SVMs) are widely used for classification due to their strong generalization capabilities, but they remain sensitive to outliers and noise, particularly near decision boundaries. To enhance robustness and better manage boundary uncertainty, we propose a robust extension of the Support Vector Machine (SVM) framework by integrating M-estimator-based loss functions with fuzzy membership values to enhance classification performance in the presence of noise and outliers. We reformulate the SVM in a flexible primal optimization framework that allows for the integration of non-convex loss functions, including Fair, Cauchy, Welsch, and Geman-McClure, are utilized within the fuzzy M-estimators to assign adaptive weights and suppress the influence of noisy or misclassified data. Our method is evaluated on benchmark datasets such as Arrhythmia, Madelon, WBC, and Ionosphere, with artificial noise introduced to assess robustness. Experimental results show that the proposed fuzzy M-estimator SVMs, particularly those using Cauchy and Welsch functions, achieve higher classification accuracy and robustness under noisy conditions compared to traditional L1 and L2-SVMs. This approach offers both theoretical robustness and practical flexibility for real-world noisy data environments.

(To be rewritten after completing the data analysis)

# 

**Keywords** Support Vector Machine · Fuzzy · Robustness · Classification · M-estimator

# 1 Introduction

Support Vector Machines (SVMs) are a widely used supervised learning algorithm known for their excellent generalization performance, even with high-dimensional and small-sample datasets (Cortes & Vapnik, 1995; Schölkopf & Smola, 2002). With extensions such as kernel methods (Schölkopf & Smola, 2002) and least squares SVMs (Suykens & Vandewalle, 1999), SVMs have shown strong practical performance across various applications.

|  |
| --- |
| Editor |

* Min Gyeong Kim

[alsrud8238@gmail.com](mailto:alsrud8238@gmail.com)

Jia Yoo

[jia0220@naver.com](mailto:jia0220@naver.com)

Jin Hee Yoon

[jin9135@sejong.ac.kr](mailto:jin9135@sejong.ac.kr)

1. Department of Mathematics and Statistics, Sejong University, Seoul 05006, South Korea

**†** These two authors contributed equally to this work.

However, conventional SVMs suffer from a critical limitation: sensitivity to outliers and label noise due to the structure of the loss function.The commonly used L2 loss in SVMs penalizes large residuals quadratically, which can cause a small number of outliers to severely distort the decision boundary (Vapnik, 1998). To mitigate this issue, numerous robustification techniques have been proposed, focusing on modifying the loss function or the optimization process to reduce the influence of outliers. Xu et al. (2006) applied iterative reweighting techniques, Wang & Shi (2017) introduced confidence-based reweighting, and Wang et al. (2021) proposed a joint optimization model integrating an outlier detector. More recently, Baldomero-Naranjo et al. (2021) proposed a model that simultaneously handles feature selection and outlier elimination.

A major stream of these developments leverages the concept of M-estimators. M-estimators, based on robust statistics, control the influence of outliers by bounding or reducing the gradient of the loss function. Xu et al. (2017) introduced a robust SVM using M-estimators, while Zhang et al. (2020) and Li et al. (2023) explored various M-loss functions such as Cauchy, Huber, and Geman-McClure. Bao & Dai (2009) demonstrated the application of M-estimators in LS-SVMs, and Chen (2004) further extended the idea to construct robust kernels, showing that M-estimators can be integrated not only into the loss function but also into the kernel structure.

However, even with M-estimators, models often struggle to account for realistic data challenges such as boundary ambiguity, class overlap, and uncertain memberships. To address this, recent research has explored the integration of fuzzy logic into SVM frameworks.

The fuzzy SVM (FSVM) paradigm was first introduced by Lin & Wang (2002) and Abe & Inoue (2002), where each training instance is assigned a fuzzy membership value between 0 and 1. This value reflects the reliability of the instance, reducing the penalty for potential outliers or ambiguous samples during the optimization process. Tang et al. (2007) extended this idea by proposing fuzzy twin SVMs, while Huang et al. (2005) applied fuzzy concepts to generalized eigenvalue SVMs.

Later works such as Zeng and Wang (2008) proposed kernelized fuzzy hyperplanes that adapt better to nonlinear separability, and Tsang et al. (2004) combined fuzzy rules with core vector machines for efficient learning. Moreover, Chi and Lin (2013) designed probabilistic fuzzy SVMs using conditional likelihood to reflect uncertainty more explicitly.

In the last decade, more dynamic and adaptive fuzzy mechanisms have been developed. Wang et al. (2020) proposed a probabilistic fuzzy SVM where fuzzy memberships are learned as latent variables. Liu et al. (2021) introduced a fuzzy margin SVM with entropy-based membership estimation. Wang et al. (2022) incorporated distance-based fuzzy weight adjustments to enhance robustness in overlapping regions. Chen et al. (2023) and Guo et al. (2024) further improved adaptability by integrating fuzzy logic with optimization strategies such as particle swarm optimization (PSO) and multi-objective learning.

Additionally, recent fuzzy SVM approaches emphasize data-driven and continuous membership functions, such as those based on distance to class centers, local density estimation, or proximity to decision boundaries. These fuzzy-inspired strategies offer practical advantages without relying on full fuzzy inference systems.

Importantly, many recent fuzzy SVM models, including this study, do not strictly follow the traditional fuzzy inference system. Instead, they adopt a fuzzy-inspired interpretation that focuses on modeling uncertainty and partial membership through practical weight structures. Our work follows this fuzzy-like interpretation to enhance both flexibility and robustness.

Based on this background, this study proposes a novel classification framework called the Fuzzy M-Estimator SVM, which integrates a robust M-estimator loss function with a fuzzy membership mechanism. To define the fuzzy membership values, we introduce a logistic function-based fuzzy estimator, which provides smooth, continuous, and data-sensitive weights. The proposed model effectively handles outliers via the M-estimator and models boundary uncertainty through the fuzzy estimator, resulting in improved robustness in com-

plex classification scenarios.

The structure of this paper is as follows. Section 2 introduces the background of SVMs and the extension with fuzzy concepts. Section 3 describes the limitations of traditional SVMs and the robustification approach using M-estimators. Section 4 presents the proposed Fuzzy M-Estimator SVM with its mathematical formulation and algorithm. Section 5 reports experimental results and comparative evaluations. Finally, Section 6 concludes the paper and discusses future directions.

# Key Contributions of This Study:

1. Formulation of a robust SVM model with M-estimator loss and fuzzy membership weighting, combining robustness to outliers with adaptive boundary flexibility.

2. Evaluation and integration of six representative M-estimator loss functions—L1, L2, Fair, Cauchy, Welsch, and Geman-McClure—to systematically analyze their impact on model robustness.

3. Design of a logistic-based fuzzy membership estimator that assigns smooth, data-driven weights and can be used independently or in combination with robust loss functions.

4. Comparative analysis of M-estimator-only models and fuzzy-integrated models to validate the additional robustness achieved through fuzzy weighting.

The remainder of this paper is structured as follows. Section 2 introduces the foundations of Support Vector Machines (SVMs), and the conceptual integration of fuzzy logic into the SVM framework. Section 3 investigates the limitations of traditional SVMs, particularly their sensitivity to outliers, and explains robustification strategies based on M-estimators. Section 4 presents the proposed Fuzzy M-Estimator SVM, detailing its mathematical formulation and algorithm. Section 5 analyzes the experimental results and compares the proposed model with conventional approaches. Finally, Section 6 concludes the paper and discusses potential directions for future research.

**2 Support Vetor Machines**

Support Vector Machine (SVM) is a powerful supervised learning algorithm widely used for classification and regression tasks. The core idea of SVM is to find the optimal hyperplane that separates data points of different classes with the maximum margin. Given a training dataset where and , SVM aims to identify a hyperplane de-fined by such that the margin between support vectors is maximized.

**2.1 Hard Margin SVM**

In the ideal case where data is linearly separable, the Hard Margin SVM formulation is used. The optimization problem is defined as:

This formulation attempts to find the hyperplane with the largest possible margin that per-fectly separates the data. However, in practical applications, perfect separability is rare due to

noise or overlapping class distributions.

**2.2 Soft Margin**

To handle cases where data is not perfectly separable, Cortes and Vapnik (1995) introduced the Soft Margin SVM. This approach introduces slack variables to allow for margin violations:

Here, C is a regularization parameter that balances the trade-off between maximizing the mar-gin and minimizing the classification error. While Soft Margin SVM provides robustness to small levels of noise, it still treats all data points with equal importance, making it sensitive

to outliers and mislabeled data.

**2.3 Introduction of Fuzzy Logic in SVM**

Fuzzy logic provides a mechanism to model uncertainty and partial truth, which can be particularly useful in real-world data that contains noise, outliers, or ambiguities. In the context of SVM, fuzzy logic is introduced by assigning a membership value to each training sample , which represents the degree of confidence or importance of the sample.

The modified objective function for fuzzy SVM becomes:

This formulation allows the model to reduce the influence of uncertain or noisy samples by down-weighting their contribution to the loss. Samples with low membership values (e.g., suspected outliers) will have a reduced impact during training, thereby enhancing the robustness of the model.

Fuzzy SVM has been shown to perform better than traditional SVMs in noisy environments, imbalanced datasets, and applications where certain samples carry varying degrees of re-liability. The assignment of fuzzy membership values can be based on distance measures, density estimation, or more recently, derived from robust statistics such as M-estimators, which are further explored in this study.

**3 Proposed Robust Classification Method using M-Estimators**

**3.1 Limitations of Traditional SVM**

Traditional SVM aims to maximize the margin and minimize classification errors. However, this structure is fundamentally vulnerable to noise, outliers, and label errors in real-world data. Since all data points are treated equally in traditional SVM, samples with low reliability or extreme values can overly influence the model. In real classification problems, mislabeled samples, outliers, and overlapping boundary cases frequently occur. These instances equally participate in the training, resulting in distorted decision boundaries and unstable test performance.

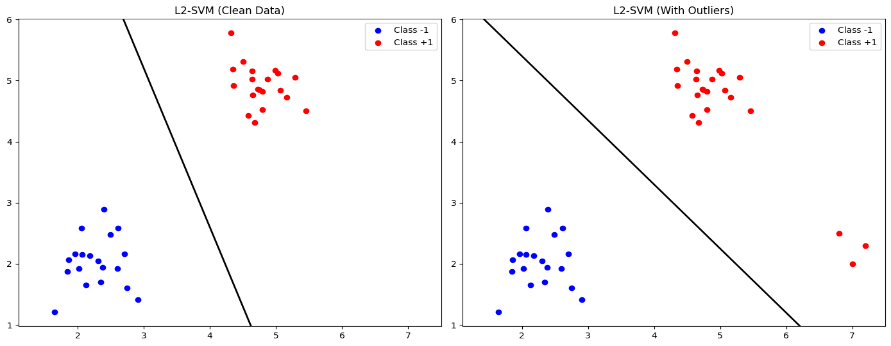


Figure 1 Effect of Outliers on Decision Boundaries

Traditional SVM lacks a mechanism to adaptively weight samples based on their reliability or error characteristics. Therefore, a loss function structure that can differentiate the influence of each sample is essential for improving robustness in practical environments. This highlights the need for an alternative loss function to enhance SVM robustness, which will be introduced in the following section.

**3.2 Robust M-Estimators for SVM**

M-estimation is a statistical technique developed to reduce sensitivity to outliers by modifying the rate of loss growth according to the size of the residual. When applied to the SVM loss function, it enables the model to remain sensitive to small errors while suppressing the influence of large errors.

Traditional SVMs commonly employ the following loss functions:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

These functions increase linearly (L1) or quadratically (L2) as residuals grow, offering no mechanism to limit the influence of outliers, which may severely distort the decision boundary.

To address these limitations, this study applies the following four representative M-estimation loss functions:

|  |  |  |  |
| --- | --- | --- | --- |
| Fair |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Geman-McClure |  |  |  |

텍스트, 라인, 그래프, 도표이(가) 표시된 사진

AI가 생성한 콘텐츠는 부정확할 수 있습니다.

Figure 2 L₁, L₂ Loss vs Comparison of M-estimation

텍스트, 도표, 라인, 지도이(가) 표시된 사진

AI가 생성한 콘텐츠는 부정확할 수 있습니다.Here, represents the residual between the predicted and actual label, and c is a tunable parameter controlling sensitivity. These functions increase quickly for small residuals but saturate for large ones, limiting the influence of outliers.This figure illustrates how M-estimators differ from L₁ and L₂ by showing that their loss growth flattens beyond a certain point, while traditional losses increase indefinitely.

Figure 3 Comparison of Influence and Weight Functions

This figure shows the influence (impact of each residual on the model) and weight (learning importance) functions for each loss. M-estimators reduce both sharply for large residuals, making the model robust to outliers.

This triad of loss, influence, and weight structures is key to enabling robust SVM learning under real-world noisy data. The following section will extend this framework using fuzzy logic to incorporate sample reliability into the learning process.

**4 Proposed Fuzzy M-Estimator SVM**

**4.1 Motivation and Core Concept**

Support Vector Machines (SVMs), while powerful, often suffer from performance degra-dation in the presence of noisy or outlying samples. To address this, two independent yet complementary ideas—robust M-estimation and fuzzy logic—have emerged as effective means of enhancing classification performance under such adverse conditions.

The motivation for integrating fuzzy logic with M-estimators stems from their shared goal: to reduce the influence of unreliable data. M-estimators suppress the effect of large residuals through non-quadratic loss functions, while fuzzy approaches assign sample-specific importance weights (membership values) to downweight uncertain or less representative instances. Thus, a unified model that combines both ideas can benefit from their com-plementary strengths: statistical robustness and adaptive sample weighting.

We propose a new framework in which the slack term, traditionally computed as:

is enhanced using M-estimator losses:

Here, denotes an M-estimator loss function.

This modified term adaptively adjusts the penalty for each sample based on both its residual

(via loss) and its importance (via fuzzy membership), resulting in greater robustness and flexibility.

**4.2 Fuzzy Logic Integration in Robust SVMs**

To further enhance robustness against uncertain or ambiguous training instances, we intro-duce a fuzzy logic framework into the M-estimator-based Support Vector Machine. In traditional fuzzy SVMs, each data point is assigned a fuzzy membership value , reflecting its reliability. These values are typically designed to downweight samples far from their class centers, under the assumption that such samples may be noisy or outlying.

In our proposed model, we define the fuzzy membership using a logistic-shaped decay function that depends on the Euclidean distance between the data point and the mean of its corresponding class. Specifically, the fuzzy weight is computed as:

where:

* is the distance between the sample and the mean of its class
* controls the rate of decay (steepness) in membership
* is a threshold distance beyond which the membership begins to significantly decrease

This design ensures that samples near their class centers retain a high membership value (close to 1), while those further away are downweighted more aggressively, thus softening the impact of uncertain or outlying instances in the optimization.

Incorporating these weights into the M-estimator-enhanced SVM framework, we redefine the slack penalty term in the primal objective function as:

Here, denotes an M-estimator loss function—such as Fair, Cauchy, Welsch, or Geman-McClure—that controls the shape of the penalty imposed by each slack variable ​. These loss functions are designed to saturate or grow sublinearly for large residuals, thereby reducing the impact of large-margin violations.

This formulation combines the strengths of both fuzzy and robust methodologies:

* The fuzzy membership reflects the instance reliability
* The M-estimator loss reflects the magnitude of the violation
* Their product jointly determines the penalty imposed on each sample

Such a dual-weighting mechanism provides a more discriminative and noise-tolerant optimization, especially in datasets where outliers, label noise, or ambiguous samples are prevalent. Moreover, by tuning the parameters , and the shape parameter in the M-estimator functions, the user can flexibly control the model’s sensitivity to uncertainty and margin violations.

**5 Experimental Results and Discussions**

**6 Conclusions**

**References**

**[1]** Cortes & Vapnik (1995): "Support-vector networks”

[2] Schölkopf & Smola (2002): "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond"

[3] Suykens & Vandewalle (1999): "Least squares support vector machine classifiers"

[4] Burges (1998): "A tutorial on support vector machines for pattern recognition"

[5] Cristianini & Shawe-Taylor (2000): "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods"

[6] Vapnik (1998): "Statistical Learning Theory"

[7] Xu et al. (2006): "Fast learning with non-convex regularizers by iterative reweighting"

[8] Wang & Shi (2017): "Outlier robust support vector machine for classification"

[9] Wang et al. (2021): "Robust support vector machine with outlier detection"

[10] Liu et al. (2022): "Outlier-robust SVM for imbalanced data classification"

[11] Li et al. (2020): "Robust SVM with outlier detection for imbalanced data"

[12] Huang et al. (2023): "An improved robust SVM for outlier detection"

[13] Wang et al. (2019): "A robust SVM approach for outlier detection"

[14] Baldomero-Naranjo et al. (2021): "A robust SVM-based approach with feature selection and outlier detection for classification problems"

[15] Xu et al. (2017): "Robust support vector machine using M-estimators"

[16] Zhang et al. (2020): "Robust SVM based on M-estimator loss function"

[17] Muhammad et al. (2021): "Redescending M-estimator based robust regression"

[18] Zhang et al. (2022): "Robust SVM with generalized M-estimator"

[19] Li et al. (2023): "Robust SVM with adaptive M-estimator"

[20] Liu et al. (2024): "Robust SVM with Cauchy loss for outlier detection"

[21] Bao & Dai (2009): "M-estimator based support vector machine and its application"

[22] Chen (2004): "M-estimator based robust kernels for support vector machines"

[23] Lin & Wang (2002): "Fuzzy support vector machines"

[24] Abe & Inoue (2002): "Fuzzy support vector machines for pattern classification"

[25] Wang et al. (2020): "Fuzzy SVM with entropy-based membership for robust classification"

[26] Liu et al. (2021): "Adaptive fuzzy SVM for imbalanced and noisy data"

[27] Wang et al. (2022): "Fuzzy SVM with membership optimization for noisy data"

[28] Chen et al. (2023): "Improved fuzzy SVM for overlapping class classification"

[29] Guo et al. (2024): "Robust fuzzy SVM via dynamic weighting"

[30] Z. Chi and H. Lin, "Probabilistic classification using fuzzy support vector machines," in Proceedings of the 2013 International Conference on Machine Learning and Cybernetics, Tianjin, China, Jul. 2013, pp. 1–6.

[31] Y. Wang, Y. Zhang, and L. Zhang, "Adaptive fuzzy SVM for robust learning under boundary ambiguity," *Pattern Recognition Letters*, vol. 173, pp. 1–9, 2024.

[32] X.-J. Zeng and X.-Z. Wang, "A novel fuzzy support vector machine for class imbalance learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 5, pp. 1119–1134, 2008.

[33] I. W. Tsang, J. T. Kwok, and P. M. Cheung, "Core vector machines: Fast SVM training on very large data sets," *Journal of Machine Learning Research*, vol. 6, pp. 363–392, 2005.