Ramp-KSVCR (Ramp Loss K-Support Vector Classification-Regression);

A sparse and robust methodology for multi-class classification problem

This is a guideline about our proposed method named “Ramp Loss K-Support Vector Classification-Regression; Ramp-KSVCR”. In summary, we proposed a precise, sparse and robust approch for multi-class classification problem based on the Ramp Loss K-Support Vector Classification-Regression. The main objectives of this research are to address the following issues;

1) Enhance the performance of K-SVCR method on some datasets that have highly imbalanced and skewed classes’ distribution,

2) Decrease the sensitivity of SVM and its extensions to the presence of noises and outliers in the training sets by introducing the Ramp loss function to our model;

3) The proposed Ramp-KSVCR model is a non-differentiable non-convex optimization problem, hence, we took Concave–Convex Procedure (CCCP) to solve this model.

To make our model more applicable in the large-scale setting and to reduce the training time, Alternating Direction Method of Multipliers (ADMM) is used to solve sub-quadratic programming problems in each iteration of our CCCP.

The details of proposed method is published in the following paper (Please cite this paper if you use this code):

Reference:

**Bamakan**, S. M. H., Wang, H., & Shi, Y. (2017). ***Ramp loss K-Support Vector Classification-Regression; A robust and sparse multi-class approach to the large-scale intrusion detection problem***. Knowledge-Based Systems, Under Review.

\*\*\*\*\* Instruction \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Ramp-KSVCR Toolbox.V4

This is a source code of “Ramp Loss K-Support Vector Classification-Regression” which is a precise, sparse and robust approch for multi-class classification problem

**Platform**: MATLAB

**Copyright**: Bamakan, S. M. H., Wang, H., & Shi, Y. (2017).

**E-mail**: S.M.Hosseini Bamakan: [s\_mojtabahossini@yahoo.com](mailto:s_mojtabahossini@yahoo.com)

Huadong Wang: [huadw2012@163.com](mailto:huadw2012@163.com)

**Researchgate page:** <https://www.researchgate.net/profile/Smojtaba_Hosseini_Bamakan>

**Github page:** <https://github.com/smhbamakan/Ramp-KSVCR>

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Data sets provided as demo:**

*Demo-Dataset.mat*: This is a random sample of NSL-KDD dataset as multi-class, highly imbalanced and skewed dataset. This dataset just provided as an example to show how our algorithm works. It should be noted that the experiments has been done in the paper are not based on this demo-dataset.

*XandY.mat*: To examine the effect of the ramp loss function on the performance of K-SVCR and to find out how this non-convex loss affects the decision boundary of the Ramp-KSVCR model, we performed some examinations with artificial data in *R2* dimension.

**Scripts:**

*Gridsearch.m:* This file can be used to do the grid search and cross-validation. The following parameters should be determined or tuned before apply the model.

obj.parameters.C ; is a constant penalty parameter(C>=0).

obj.parameters.k ; is kernel width. Here gamma as RBF kernel parameter (0,1]

obj.parameters.e ; insensitive paramter function([0,1))

obj.parameters.t ; Ramp loss function parameter((e,+Inf))

obj.parameters.s ; Ramp loss function parameter((-Inf,1])

obj.parameters.threshold=0;

obj.nOfFolds = 5 ; Number of fold for cross-validation techniques.

obj.method = Algorithm; Ramp-KSVCR is used as our method

obj.cvCriteria = ACC; Accuracy is used for cross-validation evaluation

*RampKSVCR\_Evaluation.m:* This file can be used to evaluate the performance of Ramp-KSVCR without grid search.

**It should be noted that:**

Parameter which is used to control the width of insensitive margin, is an influential parameter affect the sparsity and overall performance of the model.

We set in Algorithm 3. and are two parameters of ramp loss function which determine the effectiveness of ramp-KSVCR in dealing with the outliers. According to the ([Bottou & Giles, 2011](#_ENREF_9)), if then ; it means that moving toward the large negative value will diminish the effect of the Ramp loss on the outliers. In this research we set searching rage for and as follows: and For implementation of our model Ramp-KSVCR, we recommend that the value of *t* should be larger than and as recommended in ([Y. Wu & Liu, 2012](#_ENREF_83)), choosing , where is the number of classes.

**Plots section**

*KSVCRshow.m:* This file can be used to plot the decision boundary of K-SVCR model with artificial data in *R2* dimension.

Besides to the aforementioned parameters, there is an option to choose different plots as defined in the following:

plt = 2; It has two options; 1 and 2;

1 = just show the decision region with different colors;

2= show the decision hyperplane between class 1 and class 3.

*RampKSVCRshow.m:* This file can be used to plot the decision boundary of Ramp-KSVCR model with artificial data in *R2* dimension.

The options are as same as mentioned above.

*RampKSVCRshowEpsilon.m:* This file can be used to show the variation of accuracy with different values for .