



The Open Cybernetics & Systemics Journal

Content list available at: www.benthamopen.com/TOCSJ/

DOI: 10.2174/1874110X01610010155



RESEARCH ARTICLE

Optimization and Analysis on Fuzzy SVM for Targets Classification in Forest

Ding Xiaokang^{*,1}, Yan Lei², Yu Jianping¹ and Zhou Zhaozhong¹¹Quzhou University, Quzhou, Zhejiang, China²Beijing Forestry University, Beijing, China

Received: March 14, 2016

Revised: July 16, 2016

Accepted: July 21, 2016

Abstract: In this paper, machine learning technology was introduced to forestry machine, and harvesting targets data were classified based on fuzzy support vector machine (FSVM). Fuzzy membership function largely determines the classifier results, so a new membership function was proposed to improve classification performance. Clustering center was selected based on a K-means algorithm, then the membership function was determined *via* comparing distance between samples to the positive and negative clustering centers in feature space, respectively. It has good learning ability and generalization performance by the experiment with common SVM and representative Fuzzy-SVM. Besides, this model is applied in harvesting target detection, which improves classification accuracy and satisfies with the request of forestry machine.

Keywords: Class imbalance learning, Fuzzy Membership, K-means, SVM, Target Classification.

1. INTRODUCTION

Support vector machine (SVM) is a machine learning algorithm first proposed by Cortes and Vapnik in 1995 [1]. Compared with traditional artificial neural networks, SVM does not only simplify the learning algorithm, but also improves the technical performance, especially the generalization. Consequently, SVM has become a hot spot in the field of machine learning in recent years.

Currently, SVM algorithm has been widely used in pattern recognition, regression estimation, probability density function estimation, etc. Agrawal S. proposed a novel color image classification approach by using SVM, and the experimental results showed that this SVM-based method could improve the color image classification accuracy [2]. Kai-Ying Chen *et al.* established an SVM-based model to analyze the failure of turbines in thermal power facilities, and the results showed that SVM outperformed linear discriminant analysis (LDA) and back-propagation neural networks (BPN) in classification performance [3]. Huimin Qian *et al.* put forward a system to recognize multiple kinds of activities from videos by an SVM multi-class classifier with binary tree architecture, and the tests on a home-brewed activity dataset and the public Schuld's dataset both confirmed the perfect identification performance and high robustness of the system [4].

However, in practical applications, many outliers may make the sample data not ideal as expected, which have a significant weakening impact on the generalization performance of SVM. In view of this, Chunfu Lin and Shengde Wang put forward a fuzzy vector machine algorithm to calculate the fuzzy membership for each sample, by which the effects of noise on SVM could be eliminated effectively [5]. Based on FSVM, many kinds of estimating methods for fuzzy membership functions were presented to improve the performance of SVM. XF Jiang calculated fuzzy membership in the feature space, which is represented by kernels. It improved classification accuracy and generalization

* Address correspondence to this author at the Quzhou University, Quzhou, Zhejiang, China; Tel: +86-13967018849; E-mail: dingxiaokang1985@126.com

[6]. Rukshan Batuwita and Vasile Palade combined class imbalance learning (CIL) methods and FSVM to solve the imbalance classification problem, and the tests on many different datasets indicated that the proposed FSVM-CIL method was very effective for CIL, especially in the presence of outliers and noise in the datasets [7]. RK Sevakula used clustering algorithm to find outliers, then assigned these outliers membership value lower than “1” and the other sample with value equal to “1” [8]. YW Cheng proposed a hybrid system. They applied K-NN method to model a preprocessor to evaluate locally the contaminant level of the degenerated training data, and used FSVM as a postprocessor to learn globally an optimized classifier [9]. ON Almasi presented a method to calculate membership function and solve model selection problems based on firefly algorithm, which is a nature-inspired optimization algorithm [10].

All the above FSVM models had good performance according to the experimental results. However, they ignored an important problem that the outliers and the support vectors are all far away from the clustering center. The lower membership values for support vectors will also reduce accuracy of classifier. In order to solve this problem, in this paper, a new fuzzy membership estimating method is proposed for FSVM. The proposed model was compared with common SVM and representative Fuzzy-SVM. It was also applied to targets classification in forest. The experimental results showed that the proposed method can effectively eliminate the noise impact and improve the classification accuracy.

2. SUPPORT VECTOR MACHINE

As a new machine learning algorithm, the core idea of SVM is to map non-linear problem in the low-dimensional space to high-dimensional space so that the non-linear problem can be transformed into linearly separable problem. SVM only concerns the points nearest to the hyperplane (called as support vectors), and the hyperplane can be defined by maximizing the distances from support vectors to itself. Thus, the objective function for solving the hyperplane can be expressed as follows:

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 \\ \text{s.t.} \quad & y^{(i)} (w^T x^{(i)} + b) \geq 1, i = 1, 2, \dots, m \end{aligned} \quad (1)$$

However, outliers may make the hyperplane move easily. Considering the high sensitivity of SVM to outliers, slack variable ζ is introduced to the quadratic programming problems to permit the existence of outliers. Then, the objective function can be converted into the following form:

$$\begin{aligned} \min_{\gamma, w, b} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \\ \text{s.t.} \quad & y^{(i)} (w^T x^{(i)} + b) \geq 1 - \xi_i, i = 1, \dots, m \\ & \xi_i \geq 0, i = 1, \dots, m \end{aligned} \quad (2)$$

where penalty parameter C is used to limit the impact of outliers on the objective function. This modified model is called as soft interval classifier.

To solve the objective function, Lagrangian operator is introduced into the formula. By simplification and conversion, the hyperplane can be expressed as follows:

$$f(x) = \text{sgn} \left(\sum_{i=1}^m \alpha_i y_i K(x_i, x) + b \right) \quad (3)$$

where $K(x_i, x)$ is the kernel function for solving non-linear separable problems. In this paper, radial basis function (RBF) is selected as the kernel function, which can be expressed as follows:

$$K(x,y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{4}$$

3. FSVM AND FUZZY MEMBERSHIP FUNCTION

3.1. Fuzzy SVM

In soft interval classifier, the value of penalty parameter C cannot be too large or too small in order to ensure the effect of classifier. Hence, fuzzy membership s_i is introduced to SVM. For a training dataset $\{x_1, x_2, x_1, \dots\}$, there is a label y_i for each x_i , denoted as $\{(x_1, y_1), (x_2, y_2), \dots\}$. After s_i is introduced, the dataset is denoted as $\{(x_1, y_1, s_1), (x_2, y_2, s_2), \dots\}$, and the objective function can be rewritten as:

$$\begin{aligned} \min_{\gamma, w, b} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m s_i \xi_i \\ \text{s.t.} \quad & y^{(i)}(w^T x^{(i)} + b) \geq 1 - \xi_i, i = 1, \dots, m \\ & \xi_i \geq 0, i = 1, \dots, m \end{aligned} \tag{5}$$

where s_i takes a value between 0 and 1 ($0 < s_i \leq 1$). s_i represents the probability of each sample belonging to its label y_i . Meanwhile, it counteracts the impact of penalty parameter C on classifiers.

Therefore, an appropriate membership function is very important for FSVM model. Firstly, the lower bound of the membership should be defined. Secondly, the membership function should be constructed according to the characteristic datasets. At present, the commonly used membership calculation methods are mainly based on the distances from sample points to the class center.

3.2. Proposed Fuzzy Membership Function for FSVM

3.2.1. Clustering Center Selection

In order to calculate fuzzy membership, the class center for samples should be determined. Commonly used clustering algorithms include k-means, STING, CLIQUE and CURE. K-means algorithm is a clustering algorithm based on partition, and it is the most widely used clustering method. In this research, the class center was calculated based on a K-means algorithm proposed in Ref [11]. This method clusters samples by calculating density distribution, and its specific steps are as follows:

1. Calculate $d(x_i, x_j)$, which is distances between every two points in the sample set.
2. Calculate means of all distances.

$$Meandist(S) = \frac{1}{n(n-1)} \times \sum d(x_i, x_j) \tag{6}$$

3. For $i = 1: k$, calculate density for all samples.

$$density(x_i) = \sum_{j=1}^n u(MeanDist - d(x_i, x_j)) \tag{7}$$

$$\text{where } u(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \tag{8}$$

4. The sample with the max density value is chosen as class center.

3.2.2. Class Imbalance Learning (CIL)

Although SVMs often produce effective solutions for balanced datasets, they are sensitive to the imbalance in the datasets and produce sub-optimal models. Ref [12] has improved the standard FSVM method by combining it with the DEC (Different Error Costs) method, which is called the FSVM-CIL. In the FSVM-CIL method, the membership values for data points are assigned in such a way to satisfy the following two goals:

1. To suppress the effect of between class imbalance.
2. To reflect the within class importance of different training examples in order to suppress the effect of outliers and noise.

Our research data and most of real-world data are unbalanced, therefore the FSVM-CIL method is adopted in experiment. In the FSVM-CIL method, the membership functions are defined as follows:

$$m_i^+ = f(x_i^+)r^+ \tag{9}$$

$$m_i^- = f(x_i^-)r^- \tag{10}$$

where $f(x_i)$ generates a value between 0 and 1, which reflects the importance of x_i in its own class. The values for r^+ and r^- were assigned in order to reflect the class imbalance, such that $r^+ = 1$ and $r^- = r$, where r is the minority to majority class ratio ($r^+ > r^-$).

3.2.3. Membership Value Calculation

Many membership calculation methods didn't discriminate outliers and support vectors. In order to solve this problem, a new membership was proposed in this paper based on Ref [6], in which the membership was calculated in feature space. In the proposed FSVM model, the samples nearest the boundary of two classes were acquired firstly to distinguish support vectors and outliers. All the calculations were conducted in feature space, thus the membership can be obtained by the following formulas:

$$s_i = \begin{cases} f(d_i^+) & \text{if } \|\Phi(x_i^+) - \Phi_{cen}^+\| \geq \|\Phi(x_i^+) - \Phi_{cen}^-\| \\ f(d_{i^*}^+) & \text{if } \|\Phi(x_i^+) - \Phi_{cen}^+\| < \|\Phi(x_i^+) - \Phi_{cen}^-\| \\ f(d_i^-) & \text{if } \|\Phi(x_i^-) - \Phi_{cen}^+\| \leq \|\Phi(x_i^-) - \Phi_{cen}^-\| \\ f(d_{i^*}^-) & \text{if } \|\Phi(x_i^-) - \Phi_{cen}^+\| > \|\Phi(x_i^-) - \Phi_{cen}^-\| \end{cases}, \tag{11}$$

Where

$$f(d_i) = \begin{cases} 1 - \frac{d_i}{\max(d_i) + \beta} \\ \frac{2}{1 + \exp(\beta d_i)} \end{cases} \tag{12}$$

$$d_i = \begin{cases} \sqrt{\|\Phi(x_i^+) - \Phi_{cen}^+\|^2} \\ \sqrt{\|\Phi(x_i^-) - \Phi_{cen}^-\|^2} \end{cases}. \tag{13}$$

$$d_{i^*} = \begin{cases} \sqrt{\|\Phi(x_i^+) - \Phi_*^-\|^2} \\ \sqrt{\|\Phi(x_i^-) - \Phi_*^+\|^2} \end{cases} \tag{14}$$

Formula (12) is the linear decay and exponential decay function to calculate membership value. In the membership function (11), $\Phi(x_i^+)$ is the sample mapped to the feature space with label $y_i=1$, and $\Phi(x_i^-)$ is the one with label $y_i=-1$; $\Phi(x_{cen}^+)$ is the positive class center in feature space and $\Phi(x_{cen}^-)$ is the negative one; $\Phi(x_*^+)$ is the positive sample nearest the boundary and $\Phi(x_*^-)$ is the negative one. d_i is the distance from sample to the positive or negative class center, and d_{i^*} is the distance from sample to the sample nearest class boundary. In the feature space they can be calculated as follows:

$$\begin{aligned} \|\Phi(x_i^+) - \Phi_{cen}^+\| &= \sqrt{\|\Phi(x_i^+) - \Phi_{cen}^+\|^2} = \sqrt{\Phi(x_i^+)^2 - 2\Phi(x_i^+) \cdot \Phi_{cen}^+ + (\Phi_{cen}^+)^2} \\ &= \sqrt{K(x_i^+, x_i^+) - \frac{2}{n^+} \sum_{j=1}^{n^+} K(x_i^+, x_j^+) + \frac{1}{n^{+2}} \sum_{i=1}^{n^+} \sum_{j=1}^{n^+} K(x_i^+, x_j^+)} \end{aligned} \tag{15}$$

$$\begin{aligned} \|\Phi(x_i^+) - \Phi_{cen}^-\| &= \sqrt{\|\Phi(x_i^+) - \Phi_{cen}^-\|^2} = \sqrt{\Phi(x_i^+)^2 - 2\Phi(x_i^+) \cdot \Phi_{cen}^- + (\Phi_{cen}^-)^2} \\ &= \sqrt{K(x_i^+, x_i^+) - \frac{2}{n^-} \sum_{j=1}^{n^-} K(x_i^+, x_j^-) + \frac{1}{n^{-2}} \sum_{i=1}^{n^-} \sum_{j=1}^{n^-} K(x_i^+, x_j^-)} \end{aligned} \tag{16}$$

$$\begin{aligned} \|\Phi(x_i^-) - \Phi_{cen}^+\| &= \sqrt{\|\Phi(x_i^-) - \Phi_{cen}^+\|^2} = \sqrt{\Phi(x_i^-)^2 - 2\Phi(x_i^-) \cdot \Phi_{cen}^+ + (\Phi_{cen}^+)^2} \\ &= \sqrt{K(x_i^-, x_i^-) - \frac{2}{n^+} \sum_{j=1}^{n^+} K(x_i^-, x_j^+) + \frac{1}{n^{+2}} \sum_{i=1}^{n^+} \sum_{j=1}^{n^+} K(x_i^+, x_j^+)} \end{aligned} \tag{17}$$

$$\begin{aligned} \|\Phi(x_i^-) - \Phi_{cen}^-\| &= \sqrt{\|\Phi(x_i^-) - \Phi_{cen}^-\|^2} = \sqrt{\Phi(x_i^-)^2 - 2\Phi(x_i^-) \cdot \Phi_{cen}^- + (\Phi_{cen}^-)^2} \\ &= \sqrt{K(x_i^-, x_i^-) - \frac{2}{n^-} \sum_{j=1}^{n^-} K(x_i^-, x_j^-) + \frac{1}{n^{-2}} \sum_{i=1}^{n^-} \sum_{j=1}^{n^-} K(x_i^-, x_j^-)} \end{aligned} \tag{18}$$

where

$$K(x, y) = \langle \Phi(x), \Phi(y) \rangle = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \tag{19}$$

d_{i^*} can be obtained in the same way.

4. EXPERIMENTS

The experiments are achieved by the widely used “libsvm” toolbox under Matlab language environment. Classifier models are tested by 6 different datasets, and the geometric mean of sensitivity is used to evaluate models in order to demonstrate the comprehensive classification performance. In the results, SE represents the positive class classification accuracy, SP is the negative class classification accuracy, and $GM = \sqrt{SE \times SP}$ denotes the overall performance.

4.1. Description of Datasets

In this study, the Tree dataset is constructed by our research, and the other five are the benchmark real-world datasets obtained from UCI machine learning repository. The Tree data were obtained by multi-sensor fusion. Features of trees such as the height, width, shape, temperature, and color were extracted from 2D laser points, visible images and infrared images. The details of the six datasets are listed as Table 1. These datasets all consist of two classes and contain different sample data and attributes, covering all kinds of areas. Hence, these datasets can be considered as representatives that effectively verify the proposed FSVM models.

Table 1. Information of datasets.

Datasets	Attributes	Samples	Positive Samples	Negative Samples
Tree	6	431	381	50
Wine	13	130	59	71
Transfusion	4	748	570	178
Haberman	3	306	225	81
Pima	8	768	500	268
Quality	11	3655	1457	2198

4.2. Experimental Results of the Tree Dataset

First, Tree dataset was used to verify the proposed algorithm. Common SVM and FSVM proposed by Ref [6] were conducted in the experiment as comparisons. They were denoted by SVM and FSVM-1 respectively. By using linear decay formula and exponential decay formula, classification results were calculated with different penalty parameter C and shown in Table 2.

Table 2. Classification results of the tree dataset with different C .

Penalty Parameter C	Classification Results(%)	SVM	FSVM-1		Proposed FSVM-2	
			linear	exponential	linear	exponential
$C=2$	GM	73.96	0	73.96	0	76.60
	SE	99.45	100.00	99.45	0	99.45
	SP	55.00	0	55.00	100.00	59.00
$C=10$	GM	74.77	0	74.77	73.76	77.72
	SE	99.83	100.00	99.83	99.83	99.83
	SP	56.00	0	56.00	54.50	60.50
$C=50$	GM	76.72	38.69	76.72	79.29	78.34
	SE	99.78	99.78	99.78	99.78	99.78
	SP	59.00	15.00	59.00	63.00	61.50
$C=200$	GM	73.06	74.10	73.06	75.76	75.10
	SE	99.78	99.83	99.78	99.83	99.83
	SP	53.50	55.00	53.50	57.50	56.50
$C=500$	GM	73.08	73.76	73.08	75.76	75.10
	SE	99.83	99.83	99.83	99.83	99.83
	SP	53.50	54.50	53.50	57.50	56.50
$C=1000$	GM	75.39	75.39	75.39	75.39	75.39
	SE	99.72	99.72	99.72	99.72	99.72
	SP	57.00	57.00	57.00	57.00	57.00

In Table 2, it can be seen that when $C \leq 10$, models with linear decay formula performs not well; when $C \geq 50$, the proposed FSVM-2 with linear decay formula has the best classification accuracy. While as C increases, the results of proposed FSVM become consistent with those of other models. That is because the increment of C can weaken the effect of fuzzy membership function so as to maintain the results of the proposed FSVM consistent with other models when C is too large. It is found that the models with $C = 50$ performs best, thus the subsequent experiments are implemented with $C = 50$. Overall, the proposed FSVM-2 outperformed common SVM and FSVM-1 and it has better generalization ability.

4.3. Experimental Results of Different Datasets

In order to better verify the classification performance, the FSVM-2 models are tested by another five different benchmark datasets, and the memberships are also calculated using linear and exponential functions respectively, as shown in Table 3.

Table 3. Classification results of different datasets using linear formula.

Dataset	Classification Results(%)	SVM	FSVM-1		Proposed FSVM-2	
			linear	exponential	linear	exponential
Wine	GM	64.06	0	64.06	64.06	64.06
	SE	57.00	0	57.00	57.00	57.00
	SP	72.00	100.00	72.00	72.00	72.00
Transfusion	GM	44.36	32.31	44.40	53.36	50.02
	SE	88.14	97.86	88.29	85.43	86.29
	SP	22.33	10.67	22.33	33.33	29.00
Haberman	GM	44.88	17.25	44.94	50.30	50.75
	SE	79.00	99.20	79.20	74.40	69.60
	SP	25.50	3.00	25.50	34.00	37.00
Pima	GM	69.00	0	69.00	69.00	69.00
	SE	99.20	100.00	99.20	99.20	99.20
	SP	48.00	0	48.00	48.00	48.00
Quality	GM	71.70	59.35	71.75	71.12	72.53
	SE	66.43	41.61	66.43	70.00	68.39
	SP	77.38	84.64	77.50	72.26	76.90

From Table 3, it is obvious that the FSVM-2 has the best classification performance except the Wine and Pima datasets. According to the analysis, this discrepancy may be related with the dataset itself. In the two datasets, there is a close correlation among the data, and the class centers may overlap or approach each other. Therefore, a wholly positive or negative classification may occur according to the number of training samples in FSVM. In addition, each sample is multidimensional and has many attributes, so that it has better results when mapped to the high-dimensional space in FSVM-1 and FSVM-2. The proposed one considered the importance of support vectors and the class imbalance problem, thus it performed better than FSVM-1.

CONCLUSION

In summary, a new membership function was proposed to improve harvesting targets classification accuracy. In this method, a K-means algorithm was applied to determine the class center firstly; then decided the samples ratio to solve class imbalance problem; next, distances from sample points to the positive and negative class centers were calculated respectively, and the distances from sample points to the vectors nearest class boundary were calculated to locate support vectors; at last, fuzzy memberships were calculated by linear and exponential decay formula respectively with different distances as input. This method was applied in the feature space, with common SVM and FSVM-1 as comparison. After model construction, the performance of each model was evaluated by 6 datasets. The results showed that the proposed FSVM-2 could effectively eliminate the noise impact and improve the classification accuracy for the data with large association and many outliers. It also improved harvesting targets classification accuracy and satisfied with the request of forestry machine.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This study is financially supported by National Natural Science Foundation of China (51275272) and National Science Foundation for Young Scholars (51405263).

REFERENCES

[1] C. Cortes, and V. Vapnik, "Support vector machine", *Mach. Learn.*, vol. 20, no. 3, pp. 273-297, 1995. [http://dx.doi.org/10.1007/BF00994018]

- [2] S. Agrawal, N.K. Verma, and P. Tamrakar, "Content based color image classification using SVM", In: *Information Technology: New Generations (ITNG), 8th International Conference on IEEE*, 2011, pp. 1090-1094.
[<http://dx.doi.org/10.1109/ITNG.2011.202>]
- [3] K. Y. Chen, L. S. Chen, and M. C. Chen, "Using SVM based method for equipment fault detection in a thermal power plant", *Comput. Ind.*, vol. 62, no. 1, pp. 42-50, 2011.
[<http://dx.doi.org/10.1016/j.compind.2010.05.013>]
- [4] H. Qian, Y. Mao, and W. Xiang, "Recognition of human activities using SVM multi-class classifier", *Pattern Recogn. Lett.*, vol. 31, no. 2, pp. 100-111, 2010.
[<http://dx.doi.org/10.1016/j.patrec.2009.09.019>]
- [5] C.F. Lin, and S.D. Wang, "Fuzzy support vector machines", *Neural Netw. IEEE Trans.*, vol. 13, no. 2, pp. 464-471, 2002.
- [6] X. Jiang, Z. Yi, and J.C. Lv, "Fuzzy SVM with a new fuzzy membership function", *Neural Comput. Appl.*, vol. 15, no. 3-4, pp. 268-276, 2006.
[<http://dx.doi.org/10.1007/s00521-006-0028-z>]
- [7] R. Batuwita, and V. Palade, "FSVM-CIL: fuzzy support vector machines for class imbalance learning", *Fuzzy Syst. IEEE Trans.*, vol. 18, no. 3, pp. 558-571, 2010.
- [8] R.K. Sevakula, and N.K. Verma, "Clustering based outlier detection in fuzzy SVM", In: *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. vol. 1172-1177, 2014.
[<http://dx.doi.org/10.1109/FUZZ-IEEE.2014.6891600>]
- [9] Y.W. Cheng, T.J. Wen, and H.C. Cheng, "Distance weighted fuzzy k-NN SVM", In: *International Conference on Networking, Sensing, and Control*, IEEE, 2016.
- [10] O.N. Almasi, and M. Rouhani, "A new fuzzy membership assignment and model selection approach based on dynamic class centers for fuzzy SVM family using the firefly algorithm", *Turk. J. Electr. Eng. Comput. Sci.*, vol. 24, no. 3, pp. 1797-1814, 2016.
[<http://dx.doi.org/10.3906/elk-1310-253>]
- [11] M. Huang, Z. He, and X. Xing, "New k-means clustering center select algorithm", *Comput. Eng. App.*, vol. 47, no. 35, pp. 132-134, 2011.
- [12] R. Batuwita, and V. Palade, "FSVM-CIL: Fuzzy Support Vector Machines for Class Imbalance Learning", *IEEE Trans. Fuzzy Syst.*, vol. 18, no. 3, pp. 558-571, 2010.
[<http://dx.doi.org/10.1109/TFUZZ.2010.2042721>]

© Xiaokang et al; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International Public License (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/legalcode>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.