Serdica J. Computing **8** (2014), No 1, 15–28

Serdica
Journal of Computing

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

CLASSIFICATION OF PAINTINGS BY ARTIST, MOVEMENT, AND INDOOR SETTING USING MPEG-7 DESCRIPTOR FEATURES

Charles Welch

Abstract. Image classification is an essential problem for content based image retrieval and image processing. Visual properties can be extracted from images in the form of MPEG-7 descriptors. Statistical methods can use these properties as features and be used to derive an effective method of classifying images by evaluating a minimal number of properties used in the MPEG-7 descriptor. Classification by artist, artistic movement, and indoor/outdoor setting is examined using J48, J48 graft, best first, functional, and least absolute deviation tree algorithms. An improved accuracy of 11% in classification of artist and 17% in classification of artistic movement over previous work is achieved using functional trees. In addition classification by indoor/outdoor setting shows that the method can be applied to new categories. We present an analysis of generated decision trees that shows edge histogram information is most prominent in classification of artists and artistic movements, while scalable color information is most useful for classification of indoor/outdoor setting.

ACM Computing Classification System (1998): I.4.9, I.4.10.

Key words: Image classification, MPEG-7 Descriptors, functional classification trees, content-based image retrieval (CBIR), Categorization.

1. Introduction. Image processing and image retrieval are subjects with a growing number of applications. Content-based image retrieval (CBIR) based on specific categories is an important problem in this field [12], [13], [14]. Images can be grouped into categories based on properties of the image. Previous work suggests that MPEG-7 descriptors work well as a source of these properties [9]. MPEG-7, or "Multimedia Content Description Interface", is an ISO/IEC standard developed by the Moving Picture Experts Group [7]. This standard supports descriptions of the content of multimedia data in various ways. In the experiments described in this paper MPEG-7 descriptors for color, texture, and shape are used.

Statistical methods can be used to classify images by using these properties of MPEG-7 descriptors to construct a feature vector. Experiments presented in this paper are run using various popular decision tree algorithms. There are many advantages of using decision trees [11], however we chose to use them for the following reasons. Decision trees can efficiently be trained giving "high performance for relatively small computational effort" [11]. In addition, decision trees are human-readable. While other statistical methods produce models that are difficult for humans to understand, decision trees can be read and understood with simple Boolean logic, which makes a deeper analysis of the model possible. Finally, decision trees will tend to produce models that use few attributes. This gives us a result that can be used to quickly classify images by examining a small number of features.

In this paper we show that images from a data set of paintings can be classified into categories based on the artist, artistic movement, or whether the painting has an indoor setting. An accuracy of 11% greater for artist and 17% greater for artistic movement than previous work is shown [9], and a high accuracy is given for indoor setting classification, showing that the method can be applied to other types of categories. The next section discusses the methods used in the experiment as well as the chosen data set. Section 3 shows the results of the experiment. Section 4 presents conclusions from this study and plans for future research.

2. Method. We executed classification process in three main steps. The first step was generation of the MPEG-7 descriptor files. The second was the conversion of descriptor files into feature vectors and the last was the use of feature vectors in a decision tree algorithm. We generated descriptor files with the Multimedia Content Management System (MILOS) [1]. This produces an MPEG-7 XML descriptor file to represent each image.

The data mining software WEKA [6] was chosen to run the classification algorithms. We chose WEKA because it implements many well-known decision tree algorithms. The descriptor files are converted into feature vectors by appending comma delimited values represented in the XML into a single line. The files are then converted into rectangular arrays in ARFF format to be used by WEKA [16].

Each row of the data set corresponds to an image and the integer values representing that image's properties. All images used in this study are represented with 338 features. The received numbers of features from each of the descriptors used are as follows:

From the family of color descriptors types used include:

- *Scalable Color*—specifies a color histogram in HSV color space, encoded by a Haar transformation.
- *Color Layout*—uses the YCbCr color space with 8 bits quantization. Elements of Color Layout specify the integer arrays that hold a series of zigzag-scanned DCT coefficient values, which are derived from the corresponding component of local representative colors.
- *Color Structure*—captures both the color content and information about its spatial arrangement in a histogram. The extraction method embeds color structure information into the descriptor by taking into account all colors in a structuring element of 8x8 pixels that slides over the image.
- *Dominant Color*—extracts representative colors based on their frequency of occurrence in the image.

Two types of texture descriptors were used:

- *Edge Histogram*—describes the spatial distribution of five types of edges in sub-images, defined by dividing the image space into 4x4 non-overlapping blocks, linearized by raster scan order.
- *Homogeneous Texture*—characterizes the region texture using the energy and energy deviation in a set of frequency channels. The frequency space is partitioned with equal angles of 30 degrees in the angular direction and with an octave division in the radial direction. As a result of applying of 2D Gabor function for feature channels and consequent quantization and coding average, standard deviation, energy and energy deviation are extracted.

Lastly, one shape descriptor was used:

- *Region Shape*—is based on a set of complex 2-D transform defined on a unit disc with polar coordinates called Angular Radial Transform coefficients.

An example of part of an MPEG-7 descriptor in XML is shown in Figure 1 below.


```
<CbACCoeff2>15 14 </CbACCoeff2>
<CrACCoeff2>17 16 </CrACCoeff2>
</VisualDescriptor>
<VisualDescriptor xsi:type="ColorStructureType" colorQua
47 113 27 0 54 65 12 0 226 255 107 14 109 132 49 2 117 1
</VisualDescriptor>
<VisualDescriptor xsi:type="DominantColorType"><SpatialC
0 0 0</ColorVariance></Value><Value><Percentage>4</Perce
</Percentage><Index>160 131 93</Index><ColorVariance>0 1
1</ColorVariance></Value><Value><Percentage>1</Percentag
</Percentage><Index>101 89 68</Index><ColorVariance>0 1
1 1</ColorVariance></Value></VisualDescriptor>
```


Fig. 1. MPEG-7 XML example

The distribution of these features among the descriptor types are as follows:

– Scalable Color: 64 features;
– Color Layout: 12 features;
– Color Structure: 64 features;
– Dominant Color: 36 features;
– Edge Histogram: 80 features;
– Homogeneous Texture: 62 features;
– Region Shape: 20 features.

Several sets of feature vectors were produced with the same features but different class-labels. We manually added class labels for 120 images for each of the three types of classification. These sets are shown in Figure 2.

The chosen decision tree algorithm was then used to classify the 120 images in our data set. Cross validation techniques with a varied number of folds tell us how reliable and accurate the trees are in classifying paintings from the sets we chose. The five decision tree algorithms used in the experiments described in this paper are as follows:

- *J48*—is a univariate decision tree which is WEKA's extension of the ID3 algorithm for C4.5 decision trees [2].

- *Least Absolute Deviation Tree (LADTree)*—is a multi-class alternating decision tree. LAD uses the LogitBoost strategy [3].
- *J48 Graft*—is a grafted version of the J48 tree. For more information see [15].
- *Best First Tree (BFTree)*—builds a best-first decision tree classifier using binary split for values. More information can be found in [3], [5].
- *Functional Tree (FT)*—is a classification tree that sometimes uses logistic regression functions at certain nodes. More information can be found in [4], [10].

Renaissance	(30)	Botticelli	(15)	
		Michelangelo	(15)	indoor (60)
Baroque	(30)	Caravaggio	(15)	
		Rembrandt	(15)	
Romanticism	(30)	Friedrich	(15)	
		Goya	(15)	
Impressionism	(30)	Monet	(15)	outdoor (60)
		Pissarro	(15)	

Fig. 2. Data set distribution labels for 120 images. Number of paintings per artistic movement (left), artist (middle), and indoor/outdoor (right) are listed.

The data set chosen consists of 120 paintings. We chose this set of images as a subset of the images in [9]. These images were selected from "different web-museum sources using ArtCyclopedia as a gate to the museum-quality fine art on the Internet as well as from different Eastern public virtual art galleries and museums" [9]. The image set can be divided into movements in four equal parts. There are thirty images from the Renaissance, thirty from Baroque, thirty from Romanticism, and thirty from Impressionism. The image set can be divided into artists in eight equal parts. Fifteen images by each of these artists are included: Botticelli, Michelangelo, Caravaggio, Rembrandt, Friedrich, Goya, Monet and Pissarro. Also, sixty of the images are indoors and the other sixty are outdoors.

3. Experiments.

3.1. Comparing classification accuracy of functional trees against other tree classifiers. Experiments were done using k-fold cross-validation for 3–10 folds using various classification trees in order to maximize the correct classification percentage. The tables and figures below show percentage accuracies for J48 trees, LAD trees, J48 Graft trees, BF trees and FT trees for different cross-validation and class-labels.

Table 1 and Figure 3 show that using functional trees, an accuracy of 85.83% can be achieved with 9-fold and 10-fold cross-validation for class-label *Artist.* The functional trees always have a higher accuracy than the other tree types. Smaller learning sets give more exact results. This is because there are distinctive attributes the classifier captures for relatively small sample sizes. As the size of the sample set increases, so does the introduced noise, and the accuracy decreases.

Table 1. Percentage accuracy for decision trees for various folds, class-label: *Artist*

FOLDS	J48	LADTree	J48Graft	BFTree	FT
3	67.50	62.50	66.67	51.67	80.83
4	60.83	63.33	61.67	60.83	81.67
5	57.50	66.67	58.33	60.83	82.50
6	71.67	63.33	71.67	59.17	80.83
7	63.33	65.00	64.17	62.50	81.67
8	66.67	70.00	66.67	65.00	83.33
9	66.67	67.50	67.50	60.83	**85.83**
10	70.00	67.50	70.83	60.83	**85.83**

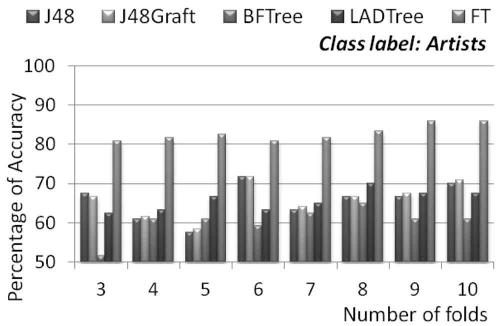

Fig. 3. Percentage accuracy for decision trees for various folds, class-label: *Artist*

Table 2 and Figure 4 show the results of classification by class-label *Movement.* Here an accuracy of 96.67% was achieved with a 10-fold cross-validation using functional trees. Again, at every fold functional trees produced a higher accuracy classification than the other types of trees. The LAD tree classifier comes close to the functional trees but never produces a higher accuracy.

Table 3 and Figure 5 show the results for class-labels *Inside/Outside.* Again, functional trees produced the highest accuracy of the classifiers listed here. It achieved an accuracy of 96.67% in a case of 4-fold cross-validation.

From this we see functional trees produced the highest accuracy of the observed tree classifiers. For the observed number of folds, classification by artist

Table 2. Percentage accuracy for decision trees for various folds, class-label: *Movement*

FOLDS	J48	LADTree	J48Graft	BFTree	FT
3	79.17	85.00	80.00	80.83	94.17
4	82.50	88.33	86.67	80.83	91.67
5	81.67	88.33	84.17	86.67	94.17
6	87.50	86.67	85.00	89.17	93.33
7	74.17	83.33	80.83	82.50	95.83
8	80.83	86.67	86.67	85.00	95.83
9	80.83	81.67	82.50	82.50	95.83
10	82.50	91.67	85.00	86.67	**96.67**

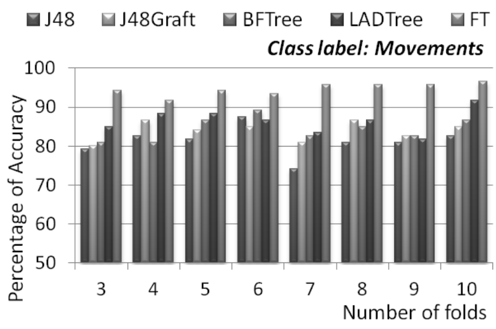

Fig. 4. Percentage accuracy for decision trees for various folds, class-label: *Movement*

Table 3. Percentage accuracy for decision trees for various folds, class-label: *Inside/Outside*

FOLDS	J48	LADTree	J48Graft	BFTree	FT
3	86.67	90.83	87.50	86.67	92.50
4	86.67	94.17	89.17	85.83	**96.67**
5	85.83	90.83	89.17	87.50	92.50
6	86.67	90.00	90.00	88.33	95.00
7	87.50	91.67	88.33	88.33	92.50
8	87.50	92.50	91.67	86.67	93.33
9	89.17	92.50	90.83	90.83	95.00
10	85.83	91.67	90.00	89.17	92.50

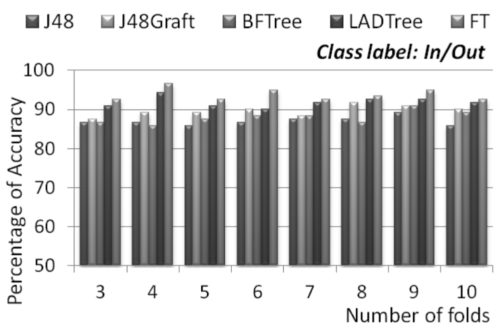

Fig. 5. Percentage accuracy for decision trees for various folds, class-label: *Inside/Outside*

and movement had the highest accuracy for 10 folds.

Classification of paintings for artist and artistic movement has been attempted previously [9]. Ivanova uses a tiling technique for preprocessing before classification. An accuracy of 75% for artist and an accuracy of about 80% are achieved for artistic movement classification. Using functional trees as shown in this paper gives an accuracy of 85.83% for artists and 96.67% for artistic movement. Also, an accuracy of 96.67% can be achieved for indoor/outdoor classification, which was not attempted by the comparison system.

3.2. Analysis of confusion matrices by functional trees. More precise analysis can be made observing produced confusion matrices. This section contains results for each case. Table 4 and Figure 6 show the confusion matrices for class-label *Artist* in case of 9-folds of the FT classifier. In the visualization of confusion matrices, the darker a square is the higher the percentage of images following in the corresponding square. We can see that some of the misclassifications are in the range of the artists that are in the same movement, for instance, Botticelli and Michelangelo from Renaissance, or Monet and Pissarro from Im-

Table 4. Confusion matrix for *Artist*, 9 folds, functional tree classifier

	Botticelli	Michelangelo	Caravaggio	Rembrandt	Friedrich	Goya	Monet	Pissarro
Botticelli	10	2	2	0	0	1	0	0
Michelangelo	1	14	0	0	0	0	0	0
Caravaggio	0	1	14	0	0	0	0	0
Rembrandt	1	0	0	14	0	0	0	0
Friedrich	1	1	0	0	12	1	0	0
Goya	0	0	0	1	0	13	1	0
Monet	0	0	0	0	1	0	13	1
Pissarro	0	0	0	0	0	0	2	13

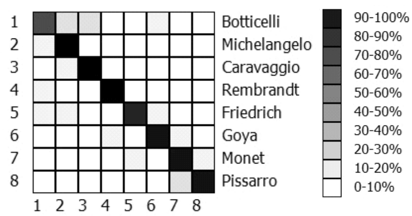

Fig. 6. Confusion matrix for *Artist*, 9 folds, functional tree classifier

Table 5. Confusion matrix for *Movement*, 10 folds, functional tree classifier

	Re nai ssan ce	Ba ro que	Ro man ti cism	Im press ion nism
Renaissance	29	0	1	0
Baroque	0	30	0	0
Romanticism	1	1	27	1
Impressionism	0	0	0	30

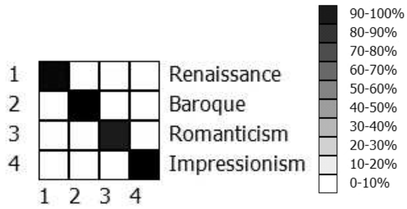

Fig. 7. Confusion matrix for *Movement*, 10 folds, functional tree classifier

pressionism. This error suggests that these paintings were classified incorrectly because the style was similar to that of an artist from the same movement.

Table 5 and Figure 7 show the confusion matrices for class-label *Movement* in case of 10-folds of the FT classifier. We can see that the color of the squares on the diagonal is in the 90–100% range showing high accuracy. Images from the Romantic period were the most misclassified. Other than Romanticism, Renaissance was the only other movement with error and this was for only one of thirty instances.

Table 6 and Figure 8 show the confusion matrices for class-label *Inside/Outside* in the case of 4 folds of the FT classifier. We can see that all instances of inside are classified correctly and that only four instances of outside are classified as inside.

Table 6. Confusion matrix for "IN/OUT", 4 folds, functional tree classifier

	In	Out
In	60	0
Out	4	56

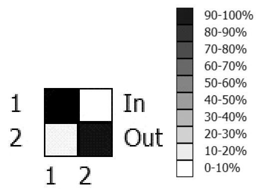

Fig. 8. Visualization of confusion matrix for *Inside/Outside*, 4 folds, functional tree classifier

3.3. Analysis of used attributes in recognition model by functional trees. For artist classification there are 73 features used by the generated functional tree for eight classes. Of these features seven of them are from the

Homogeneous Texture Type. Thirty one of the features use the Edge Histogram Type. Eleven features use the Color Structure Type. Eighteen of the features use the Scalable Color Type. The other six features are Color Layout Type attributes. The greatest portion of these five descriptors that are used is the Edge Histogram Type. 42% of the attributes use this descriptor. For artists this means that the most useful of the types of information for classification is information about the spatial distribution of various types of edges.

For movement classification there are 43 features used by the generated functional tree for four classes. Of these features, fourteen use the Scalable Color Type descriptor, ten use the Color Structure Type descriptor, sixteen use the Edge Histogram Type, two use the Color Layout Type, and one uses the Homogeneous Texture Type. This means that 33% of the features use the Scalable Color Type and 37% use the Edge Histogram Type. This shows that these are the two most useful descriptors for classifying paintings by movement.

For inside versus outside classification there are fourteen features used by the generated functional tree for two classes. Seven of the features are Scalable Color Types, another six are Edge Histogram Types and the last one is a Color Layout Type. This means that the information to use for classifying paintings for the criteria of inside versus outside are the histograms of colors used in the image, and edges distributed in the image. The spatial information about colors in the Color Layout Type only represents about 7% of the checked attributes.

Looking at these results we see that the two most useful descriptors are the Scalable Color Type and the Edge Histogram Type. These are the most used by the functional trees regardless of the type of classification.

Some examples of classified images for the three types of classification are shown in Figures 9–14.

4. Conclusion. Image classification can be achieved with high accuracy

Botticelli, *Annunciation, 1489-90* | **Michelangelo**, *Joel, 1509* | **Monet**, *Garden at Giverny, 1900* | **Pissarro**, *Landscape at Chaponval, 1880*

Fig. 9. Classification Examples for *Artist*—correct classification

Botticelli, *Annunciation, 1481*	Michelangelo, *Ignudo, 1509*	Monet, *The Water Lilly Pond, 1900*	Pissarro, *Hyde Park, London, 1890*
Botticelli classified as **Michelangelo**	**Michelangelo** classified as **Botticelli**	**Monet** classified as **Pissarro**	**Pissarro** classified as **Monet**

Fig. 10. Classification Examples for *Artist*—misclassification

Botticelli, *The Birth of Venus, 1486*	Caravaggio, *Cupid, 1601*	Goya, *The Kite, 1777*	Pissarro, *Spring in Louveciennes, 1870*
Renaissance	**Baroque**	**Romanticism**	**Impressionism**

Fig. 11. Classification Examples for *Movement*—correct classification

Botticelli, *Madonna of the Sea, 1477*	Friedrich, *Evening, 1824*	Friedrich, *Trees in the Moonlight, 1824*	Friedrich, *Mountain in Riesengebirge, 1835*
Renaissance classified as **Romanticism**	**Romanticism** classified as **Baroque**	**Romanticism** classified as **Renaissance**	**Romanticism** classified as **Impressionism**

Fig. 12. Classification Examples for *Movement*—misclassification

Botticelli, *Madonna del Libro, 1483* — **IN**

Caravaggio, *St Matthew and the Angel, 1603* — **IN**

Pissarro, *Resting in the Woods at Pontoise, 1878* — **OUT**

Monet, *The White Water Lilies, 1899* — **OUT**

Fig. 13. Classification Examples for *Inside/Outside*—correct classification

Friedrich, *Trees in the Moonlight, 1824*

Friedrich, *Greifswald Harbour, 1818–20*

Goya, *Card Players, 1777–78*

Goya, *Aquelarre, 1798*

Fig. 14. Classification Examples for *Inside/Outside*—*Outside* misclassified as *Inside*

by using functional decision trees, which is an improvement over previous work of 11% for classification of artist and 17% for classification of artist movement. In addition, a high accuracy of 96.67% for classification of indoor versus outdoor setting was shown, a task that has not previously been examined. This shows that our method of classification can be applied to new categories. Attributes from MPEG-7 descriptors can be used to represent images as feature vectors to be used in the classification process. Decision trees make decisions using a small number of calculations based on the properties in MPEG-7 descriptors. The trees give a human readable description of which features are used and how they are used to classify paintings.

Accurate classification of three kinds was shown in this paper. Classification was successfully completed by artist, movement and inside versus outside. The functional trees most used MPEG-7 attributes were the Scalable Color Type and the Edge Histogram Type. This gives insight into what information is needed to know how to classify images in various ways. Increased accuracy in classification means that image retrieval methods can be improved. The inclusion of classification by artist, artistic movement, and additionally indoor/outdoor shows that new types of image retrieval can be accomplished with similarly high accuracies. As users desire new ways to search for images this method can be applied.

Future work will expand the variety of images included in the evaluation. Applying this method to more types of images will show that it works for other data sets besides paintings. A comparison with different types of statistical methods for classification could show higher accuracy under different conditions. Also, with the knowledge of which descriptor attributes were most useful as features in classification, the generation of MPEG-7 descriptors could be optimized. Using fewer features will increase the efficiency of classification. An information retrieval system could be developed to test exactly how retrieval of images is affected by increased accuracy in classification of images.

REFERENCES

[1] Amato G., C. Gennaro, F. Rabitti, P. Savino. Milos: a multimedia content management system for digital library applications. In: Research and Advanced Technology for Digital Libraries, LNCS, Vol. **3232**, Springer Berlin/Heidelberg, 2004, 14–25.

[2] Bhargava D. N., G. Sharma, R. Bhargava, M. Mathuria. Decision Tree Analysis on J48 algorithm for Data Mining. *International Journal of Advanced Research in Computer Science and Software Engineering*, **3** (2013).

[3] Friedman J., T. Hastie, R. Tibshirani. Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, **28** (2000), No. 2, 337–407. `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.51.9525`.

[4] Gama J. Functional trees. *Machine Learning*, **55** (2004), No. 3, 219–250.

[5] Haijian S. Best-first decision tree learning. Doctoral dissertation, Ph. D. dissertation, Univ. Waikato, Hamilton, New Zealand, 2007.

[6] Hall M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, **11** (2009), No. 1, 10–18.

[7] International Standard ISO/IEC 15938-3 Multimedia Content Description Interface—Part 3. `http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34230`

[8] Ivanova K. A Novel Method for Content-Based Image Retrieval in Art Image Collections Utilizing Color Semantics. Doctoral dissertation, PhD Thesis, Hasselt University, Belgium, 2011.

[9] Ivanova K., P. Stanchev, E. Velikova, K. Vanhoof, B. Depaire, R. Kannan, K. Markov. Features for art painting classification based on vector quantization of MPEG-7 descriptors. In: Data Engineering and Management, LNCS, Vol. **6411**, Springer Berlin Heidelberg, 2012, 146–153.

[10] Landwehr N., M. Hall, E. Frank. Logistic model trees. *Machine Learning*, **59** (2005), No. 1–2, 161–205.

[11] Rokach L. *Data mining with decision trees: theory and applications.* Vol. 69. World scientific, 2008.

[12] Smith J., Ch.-Sh. Li. Image classification and querying using composite region templates. *Computer Vision and Image Understanding*, **75** (1999), No. 1–2, 165–174.

[13] Vailaya A., M. A. T. Figueiredo, A. K. Jain, H.-J. Zhang. Image classification for content-based indexing. *IEEE Transactions on Image Processing*, **10** (2001), No. 1, 117–130.

[14] Veltkamp R. C., M. Tanase. Content-based image retrieval systems: A survey. Rapport no UU-CS-2000-34, 2000.

[15] Webb G. I. Decision tree grafting from the all-tests-but-one partition. *IJCAI,* **2** (1999), 702–707.

[16] Witten I., E. Frank. Data Mining: Practical Machine Learning Tools and Techniques. 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

Charles Welch
University of Michigan, Ann Arbor
MI, 48109, USA
e-mail: `cfwelch@umich.edu`

Received January 26, 2014
Final Accepted August 25, 2014