KEER2010, PARIS | MARCH 2-4 2010
INTERNATIONAL CONFERENCE ON KANSEI ENGINEERING AND EMOTION RESEARCH 2010

ASSOCIATING COLOR WITH EMOTIONS BASED ON SOCIAL TAGGING

Haifeng FENG[a*], Marie-Jeanne LESOT[a], Marcin DETYNIECKI[a]

[a] *Université Pierre et Marie Curie - Paris 6, CNRS, UMR7606, LIP6, France*

ABSTRACT

Color is an important component to express emotions in images. The relationships between color and emotions have long been studied, both from artistic and psychological points of view. In this paper, we study the links that can be established in an unconstrained experimental setup, using a social tagging system that enables users to freely label images, without imposing any restriction regarding the label or the image choices. After collecting images labeled with emotion terms from the FlickR system, we consider both an objective color description of the images in the HSV color space, as well as a subjective semantic description, using images presenting color terms tags. The latter thus applies to images for which annotators underlined the importance of the chromatic content and leads to the identification of emotion characterization in terms of colors and reciprocally of colors in terms of emotions, although not in bijective relations.

Keywords: *emotion detection, image processing, affective computing, machine learning*

1. INTRODUCTION

The relationships between colors and emotions have long been studied, both from a theoretical point of view, formalizing artistic theories (see for instance [1]), and from a psychological point of view (see for instance [2, 3, 4]). More recently, many works have attempted to link images and emotions using machine learning tools, such as artificial neuron networks [5], fuzzy c-means [6] or support vector machines [7]. All these works are based on rather small number of examples, since one of the main problems is the constitution of an appropriate data set, containing images annotated with the studied emotions.

---

* **Corresponding author**: DAPA, LIP6, 104 avenue du President Kennedy, 75016 Paris, France, haifeng.feng@lip6.fr.

In this paper, we propose to identify relations between colors and emotions based on data extracted from a photo sharing website (www.flickr.com). Such a website is particularly interesting since it provides social tagging possibilities, i.e. it lets users tag photos they upload with freely chosen labels. Thus, millions of users tag hundreds of millions of photos, which makes it possible to construct databases with sufficient numbers of images.

The framework we propose consists of three steps: first, we build data sets by retrieving images associated to a set of emotional tags; second we extract color descriptors from the images, third we identify and study the relations between colors and emotions, using statistical tools and machine learning algorithms.

Regarding the emotions we consider, we exploit the well-known Plutchik's emotion model [8]. The latter belongs to the framework of categorical emotion model and it is made of 8 basic emotions, for each of which 3 intensity levels are defined. In this paper, we adopt the middle intensity level, thus we consider the following 8 emotions: anger, anticipation, joy, trust, fear, surprise, sadness and disgust.

Regarding the color description we successively consider two representations: an objective numerical one through the HSV color space, and a subjective linguistic one, based on color terms used to label the images. To that aim we carry out two experiments: in a first one, we download images tagged with one of the emotional terms mentioned above and extract their chromatic content in the HSV space. We then apply the decision tree algorithm C4.5 [9] to identify color-emotion relations. The obtained results underline that in such a data set, emotions associated to images depend on the interpretation of their content (e.g. a person smiling), rather than their visual appearance (e.g. dark or saturated colors).

In a second experiment, we consider a linguistic description of the image chromatic content, based on images tagged with names of colors, in addition to the emotion names. Thus we focus on images for which users considered the visual appearance of the image (in this case the chromatic description) an important feature. We then study the associations between colors and emotions, highlighting the most frequent ones.

The paper is organized as follows: after a brief introduction about social tagging and the FlickR website, Section 2 describes how the latter is used to build the image corpuses. Section 3 presents the experiment on the chromatic characterization of emotions based on a numerical color description in the HSV space. Section 4 describes the experiment based on the linguistic color description through color name tags, as well as a study of the use of emotion and color tags in FlickR. Finally, the conclusions obtained from these experiments and exploitations are stated in the last section along with the direction of our future works.

2. CORPUS CONSTRUCTION

2.1. Social tagging systems - FlickR

Many social tagging systems emerged along with the development of the web, for different types of resources such as restaurants, bookmarks in the case of the social bookmarking web service del.icio.us, or more frequently photos and videos, as is the case of the image and video hosting website FlickR. These systems allow users to share tags for supported resources: the latter are thus linked by the tags, enabling users to efficiently mine them.

In this paper, we consider the FlickR system (www.flickr.com), which is very popular in the field of photo sharing. It offers the possibility to store and share photos with family, friends and to tag them easily. Unlike some other social systems, FlickR is not a hierarchical system and it does not restrict preferences of users to tag their resources: it does not impose taxonomy for choosing the labels, which could bias the tagging step, restrict users and prevent them from finding a suitable term to express their feelings. On the contrary, FlickR makes it possible to use any tag to label the resources. This enables to construct a reliable corpus.

2.2. Retrieval of images from FlickR

In order to build an emotional image corpus, we use the web-service API and propose to use configurable queries of the form:

emotion terms [and color terms] [*restriction] [-restriction]

The first part contains emotion terms, it can be constituted of one or several terms of the 8 emotions defined in Plutchik's model. It is then optionally followed by the color terms which are represented by the 11 basic color terms, as they were defined in the reference work on color naming [10], namely black, white, red, green, yellow, blue, brown, orange, pink, purple and gray. Lastly, we add some optional restrictions, of two different types: the restriction starting with the star symbol expresses a restriction about time, i.e. the period during which images are collected. It is specified by the start and the end date. The restriction starting with "-" is used to exclude specific terms in the image labels, and filter out some images. It aims at reducing the number of images for which the semantic content determines the emotion and dominates the chromatic content. Indeed, we observed for instance that many images labeled with *anger* represent demonstrations, and cannot be related to any specific colors: their color content cannot be associated to emotions. The semantic restriction filters out such images by excluding these taboo tags. The following query illustrates the type of considered queries:

(anger OR anticipation OR … OR joy) AND (red OR green OR blue … OR yellow) *1/6/2006 *11/30/2006 –demonstration

This example extracts images tagged with at least one emotion tag and at least one color tag, uploaded between 01/06/2006 and 06/30/2009 and not labeled with the word *demonstration*. Figure 1 illustrates some images extracted by this query.

(a) (b) (c)

Figure 1: Examples of some collected images with their chromatic and emotional labels: (a) red – anger, (b) red – anger, (c) green - trust

3. EXPERIMENTS BASED ON AUTOMATIC COLOR DESCRIPTION

In this section, we describe experiments carried out to discover possible relations between color representation of image and the associated emotion tag, based on an objective numerical color description in the HSV color space.

3.1. Protocol

According to the requirements of the experiment, we construct an ensemble of images such that, first, each image is tagged with only one emotion tag and, second, the numbers of images for each emotion are balanced. We retrieve such a corpus using one query per emotion, *anger, anticipation, … joy*, and then eliminating images associated with more than one emotion term as well as those possessing the taboo tags. Finally, we take first the 1000 images in each query result to balance the number of images for each emotion.

To produce color representation of these images, we code them as histograms in the HSV color space. We divide it into 36 non-uniform subspaces [11], quantizing hue into 7 non-uniform colors, and dividing saturation and value into 3 non-uniform parts respectively.

We carry out three experiments with this data set, as described below: (i) a classification task to distinguish between the 8 emotion classes from the HSV histograms, (ii) 28 binary classification tasks, for all emotion pairs, and (iii) a binary classification task to discriminate positive emotions from negative emotions. We use the decision tree algorithm C4.5 integrated within the machine learning platform WEKA [12].

3.2. Automatic prediction of any emotions

In this experiment, we consider a classification task in 8 classes defined as the considered 8 emotions from the numerical HSV description of the image chromatic content. We assemble all images of 8 emotions in one corpus and use the C4.5 algorithm with 10 folds cross validation. We observe that the good classification rate is 20%, which is only slightly above the baseline model that randomly predicts one of the 8 classes, and that would get 12.5%. The confusion matrix of Table 1 gives details about this global value.

Table 1: Classification of 8 emotions

classified as ->	anger	anticipation	disgust	fear	sadness	joy	surprise	trust	total	entropy
anger	187	127	115	155	137	92	102	80	995	0.984
anticipation	146	136	124	130	125	124	102	102	989	0.997
disgust	124	117	235	115	93	101	125	79	989	0.973
fear	164	140	102	180	130	90	95	85	986	0.983
sadness	137	144	110	157	164	83	81	113	989	0.986
joy	128	149	89	93	90	229	123	96	997	0.973
surprise	125	148	130	108	90	123	182	92	998	0.988
trust	109	128	100	106	114	89	85	269	1000	0.961
total	1120	1089	1005	1044	943	931	895	916	7943	/
entropy	0.993	0.999	0.976	0.989	0.99	0.97	0.98	0.95	/	

The diagonal of the matrix indicates the number of images that are correctly assigned to their real emotion, for each predicted class. It can be observed that, except for *anticipation*, they correspond to the maximal value in each column. This means that the predicted class *anticipation* is not dominated by images belonging to the actual class *anticipation*, thus it does not characterize the actual class *anticipation*. On the contrary, the predicted class *trust* presents the highest number of correctly identified images, which suggests that it may be the easiest emotion to detect. Following this way, the emotions *joy* and *disgust* are the second and the third easiest ones to distinguish.

To measure the ambiguity of the obtained classes, we compute their entropy representing their purity: it quantifies their dispersion on the actual classes and can indicate if there exists a dominant emotion in the predicted class. It can be observed that the entropy values are very close one to another and very high, as all of them are above 0.9. Consistently with the previous observations, the predicted class *trust* has a slightly smaller value, while *anticipation* has the highest one.

Reciprocally, we compute the entropy of the emotions on the predicted classes (see last column of the table) to assess the dispersion of the real emotions on the predicted classes: *anticipation* is the emotion that is the most dispersed on the obtained classes, confirming the difficulty of its prediction.

3.3. Pair of emotions

Since it appears to be difficult to distinguish all 8 emotions at one time, we shift our attention to pairs of emotion: we consider 28 classification tasks in 2 classes. To that aim, we first make 28 new corpuses where each corpus includes images of only two emotions. In this experiment, we want to know if one emotion can be distinguished easily from another in each pair of emotions. As previously, we use C4.5 to test these 28 corpuses, the obtained results are shown in Table 2, ordered by the rate of good classification.

Table 2: Classification of pairs of emotions

Emotion 1	Emotion 2	Rate	Emotion 1	Emotion 2	Rate
anger	trust	65.01%	disgust	joy	60.02%
sadness	surprise	64.72%	sadness	disgust	59.91%
surprise	trust	64.66%	anticipation	sadness	57.84%
sadness	joy	64.35%	disgust	surprise	57.57%
fear	trust	63.85%	anger	disgust	57.56%
joy	trust	63.50%	surprise	joy	57.29%
fear	joy	63.39%	anticipation	joy	56.80%
sadness	trust	63.05%	fear	anticipation	56.30%
anger	joy	62.95%	surprise	anticipation	55.71%
disgust	trust	62.39%	anger	anticipation	55.49%
fear	surprise	61.69%	fear	sadness	55.49%
fear	disgust	61.37%	disgust	anticipation	54.55%
anger	surprise	61.36%	anger	fear	54.32%
anticipation	trust	61.34%	sadness	anger	53.88%

From the table, it appears that the easiest emotions to distinguish are *anger* and *trust*, whereas the most difficult are *sadness* and *anger*. Furthermore, it can be observed that the *7* experiments that include *trust* are situated in the first half part of the table, which confirms that *trust* is the easiest emotion to predict from the considered images. Inversely, emotion pairs involving *anticipation* are located to the end of the table, showing that it is difficult to be distinguished from other emotions. The good classification rate for the pair (*anticipation*, *trust*) is the lowest one in the *7* experiments involving *trust*. Still, the good classification rates remain rather low.

3.4. Positive and negative emotions

In this subsection, we consider another experiment that does not aim at distinguishing all emotions individually, but at categorizing the images as associated to positive or negative emotions, which can be associated to the task of valence prediction. We assign the 8 emotions to two principal categories: the first one contains the positive emotions, namely *surprise*, *anticipation*, *trust* and *joy;* the second one contains the negative emotions, i.e. *anger*, *disgust*, *sadness* and *fear*. After this reorganization, we consider a new classification task in these two classes. The obtained results are shown in the confusion matrix of Table 3.

Table 3: Confusion matrix

Classified as ->	Positive emotions	Negative emotions
Positive emotions	2157	1827
Negative emotions	1491	2468

The rate of good classification is 58%, which remains low: there is almost no improvement related to the performance despite this reorganization and task simplification.

One reason explaining these results may be the fact that for the images labeled with emotional tags in FlickR, the emotion is related to their semantic content, and not to their chromatic content: the filtering performed by the semantic restriction may be insufficient.

Thus the prediction of any emotional information (emotion itself or its valence) from the numerical description of the chromatic content in the HSV space has a high difficulty level: it appears to be as difficult as the detection of the object on images and the recognition of the image content.

4. EXPERIMENTS BASED ON COLOR TAGS

In this section, we consider a linguistic representation of the image chromatic content, based on images whose tags contain names of colors. We thus focus on FlickR images for which the chromatic content is explicitly indicated as important, which may reduce the bias of the image semantic content. Indeed, if an image is labeled with a color term by a user, it implies the user wants to accentuate the color role in this image, which may thus be used to predict the emotion. We propose to directly use the color tags of the images as their color description instead of their HSV histograms.

To implement this objective, we need images simultaneously labeled with at least one emotion tag and one color tag. In order to build such a corpus, we apply the following procedure: first we collect images labeled by at least one emotion tag; second we reserve images from the first step whose labels contain at least one color term; lastly, we eliminate images whose only color tags are *black* and *white*. Indeed, these two tags indicate a technical property of the images that are out of our studies, as we aim at using colors to categorize emotion.

In these queries, we still use the emotion terms defined in Plutchik's color model and the 11 basic color terms in English *black*, *white*, *red*, *green*, *yellow*, *blue*, *brown*, *orange*, *pink*, *purple* and *gray* that were defined in the color naming reference [10]. We obtain a corpus containing 22556 images during three years from 2006 to 2009.

We then perform three studies in this section: we first consider images from the emotion point of view, to discover relationships between emotions, and study the use of emotion terms in social tagging, before turning to the color point of view. Lastly, in Section 4.3, we study the association between emotions and colors.

4.1. Statistic of the association inter-emotion

The objective of the study presented in this subsection is to discover relations that may exist between emotions, for instance to determine which emotion is more frequently associated with another one.

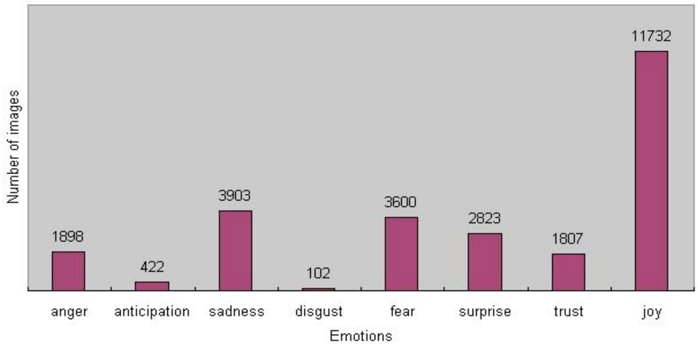

Figure 2: Distribution of emotion tags

First, Figure 2 shows the number of images for each emotion tag, and highlights that the distribution about 8 emotions is very unbalanced, and in particular underlines the overwhelming quantity of *joy* images. This may be the sign that people prefer to share their "joy" photos in website. On the contrary, the emotions *disgust* and *anticipation* only have 102 and 422 images in the ensemble of images, and are minority.

Table 4: Emotion co-occurrences

	anger	anticipation	sadness	disgust	fear	surprise	trust	joy	total
anger	0	1	75	5	69	2	2	13	167
anticipation	1	0	6	0	7	5	1	19	39
sadness	75	6	0	1	42	2	2	52	180
disgust	5	0	1	0	2	1	1	1	11
fear	69	7	42	2	0	30	12	12	174
surprise	2	5	2	1	30	0	2	43	85
trust	2	1	2	1	12	2	0	35	55
joy	13	19	52	1	12	43	35	0	175

Table 4 then shows the emotion co-occurrences among images labeled with several emotions. It must be underlined that the latter are only 506 images, i.e. about 2% of the corpus. The relative distribution of the emotions among the images that possess several emotion tags appears to be close to the overall distribution. Regarding the emotion co-occurrences, it can be observed that the emotions most related to *anger* are *sadness* and *fear*. The emotion *sadness* is more related to *anger* and *fear*, indicating that the three of them build a group of frequently co-occurring emotions. Another frequent relation is that between *trust* and *joy*. These results are consistent with intuition, however it is also to be observed that there are many images simultaneously tagged with the 2 opposite tags *sadness* and *joy*.

4.2. Statistic of the association inter-colors

We then study the relations between colors in the same manner as for emotions: Figure 3 shows the distribution of color tags among all images. It can be observed that it is unbalanced, two categories of color terms can be distinguished: *black*, *white*, *red*, *green*, *yellow* and *blue* are the frequently used terms, whereas *brown*, *orange*, *pink*, *purple* and *gray* are much rarer.

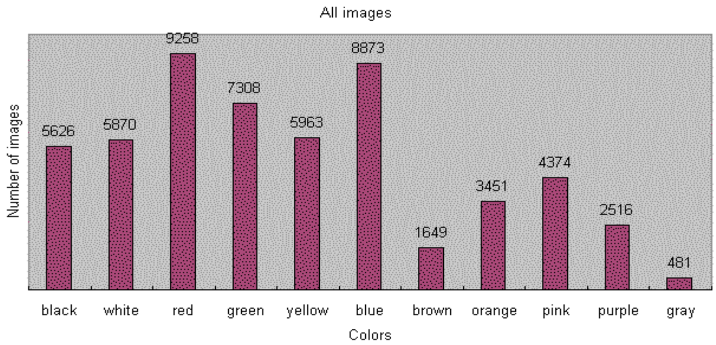

Figure 3: Distribution of color tags

It appears that 6827 images from the total corpus containing 22556 images, i.e. 30% of the images, are tagged by several color terms, whereas only 2% are labeled with several emotions: this may indicate that the emotional content of images is mostly pure and simple, while their chromatic content requires more labels. On the other hand, for 70% of the images, users use a single color term, which remains a high value. It must be underlined that they have no instructions telling them to describe the chromatic content of the image, the labeling step is fully free, and so conclusions regarding the presence of a single dominant color in the image cannot be drawn. Still, it appears that when labeling images with color words, users intuitively mostly settle for a single color.

Table 5 contains the color co-occurrences for the images labeled with several color words. It shows that for black, the most frequently co-occurring colors are red, white and blue in order of importance; for white, they are red, blue and black; for red, they are blue and green.

Table 5: Color co-occurrences

	black	white	red	green	yellow	blue	brown	orange	pink	purple	gray	total
black	0	691	861	400	430	658	186	222	207	143	105	3903
white	691	0	795	623	368	769	200	195	318	152	93	4204
red	861	795	0	1007	984	1091	182	535	485	300	22	6262
green	400	623	1007	0	772	1031	198	368	498	314	25	5236
yellow	430	368	984	772	0	851	121	548	371	286	16	4747
blue	658	769	1091	1031	851	0	192	404	498	337	46	5877
brown	186	200	182	198	121	192	0	72	57	22	13	1243
orange	222	195	535	368	548	404	72	0	245	184	15	2788
pink	207	318	485	498	371	498	57	245	0	298	12	2989
purple	143	152	300	314	286	337	22	184	298	0	4	2040
gray	105	93	22	25	16	46	13	15	12	4	0	351

4.3. Statistic in the relationships between color and emotion tags

Lastly, we examine the relations between emotions and colors, counting their co-occurrences in the images labels. When computing this table, we consider that if an image is tagged with *joy*, *yellow* and *red* for instance, then 2 co-occurrences are taken into account, namely (*joy*, *yellow*) and (*joy*, *red*).

Table 6: Emotion-color co-occurrences

	anger	anticipation	sadness	disgust	fear	surprise	trust	joy	total	entropy
black	341	39	646	15	875	291	214	1211	3632	0.807
white	227	73	455	9	419	422	238	2020	3863	0.722
red	603	75	693	17	753	648	419	2959	6167	0.756
green	215	61	416	15	465	500	436	2571	4679	0.697
yellow	112	50	278	10	300	335	166	2084	3333	0.621
blue	277	87	857	12	628	554	355	3290	6060	0.694
brown	58	22	160	6	137	145	71	425	1024	0.799
orange	89	29	172	4	183	202	103	1063	1845	0.671
pink	76	48	214	8	164	435	113	1733	2774	0.608
purple	52	20	107	4	99	140	50	757	1229	0.631
gray	14	8	121	4	72	34	9	65	327	0.787
total	2064	512	4119	104	4095	3706	2174	18159	34933	/
entropy	0.854	0.938	0.911	0.95	0.89	0.929	0.885	0.911	/	/

As shown in Table 6, in particular from the entropy values, it appears that each color has a dispersed distribution on emotions and that likewise each emotion is dispersed on the colors. In addition, most distributions about emotions are more dispersed than those of colors.

Nevertheless, regarding the color distribution, it appears that most of the *gray* images are related to *sadness*, so *gray* may be said to be mostly a *sadness* color. The same phenomenon happens in *purple* images which are mostly related to *joy*; likewise, *blue* is more associated with *sadness* and *joy*; *black* is more related to *sadness*, *fear* and, perhaps surprisingly, to *joy*.

Analyzing the results by emotions it can be noticed that from the *anger* column, 603 images are also labeled with *red*, which maximizes the co-occurrence value, reaching a proportion of 30%: this implies that the *anger* emotion is more related to *red*. Similarly, *fear* is associated with *black*, *red* and *blue*.

These preliminary analyses show that each color represents several emotions, possibly with various representative degrees, and that conversely emotions can take several colors. Other characterizations to be conducted will aim at refining these observations as well as identifying some negative associations, i.e. looking for significant color absences, instead of color presences.

5. CONCLUSION

In this paper, we considered the problem of characterizing emotions in terms of colors, in an unconstrained experimental setup, using a social tagging system that enables users to freely label images, without imposing restrictions regarding the label or the image choices. We studied the use and the distribution of both emotional and color tags, analyzing the co-occurrence matrices and performing several classification tasks to qualify the relations between colors and emotions. We took into account two types of color description: an objective, numerical one that represents images in the HSV color space and a subjective, linguistic one that exploits the color words labels used by the annotators. When using the numerical color description, the prediction task appears to have a high difficulty level, comparable to that of the automatic recognition of the image content: it would be interesting to study the relation between some objects and emotions. Still the analyses carried out on these data collected from FlickR show promising results, both for the association of emotions to colors and for the reciprocal association of colors to emotions. Future works aim at applying machine learning techniques to these data, and in particular association rules methods, to extract more precise chromatic characterizations of emotions in a social tagging framework.

REFERENCES

[1]. J. Itten, The art of color; the subjective experience and objective rationale of color. *Wiley*, 1997, translated from the German version published in 1961.

[2]. M. Hemphill, A note on adults' color-emotion associations, *Journal of Genetic Psychology*, 54, pages 275-281, 1996.

[3]. L.-C. Ou, M. R. Luo, A. Woodcock, A. Wright, A study of colour emotion and colour preference, *Color Research and Application*, 29, pages 232-240, 2004

[4]. T. Clarke, A. Costall, The emotional connotations of color: a qualitative investigation, *Color Research and Application*, Vol. 33, pp 406-410, 2008

[5]. T. Hayashi, M. Hagiwara, Image query by impression words - the IQI system, *IEEE Transactions on Consumer Electronics*, Vol. 44, pp. 347-352, 1998

[6]. K. Wei, B. He, T. Zhang, W. He, Image Emotional Classification Based on Color Semantic Description, *Lecture Notes in Computer Science (LNCS)*, Vol. 5139/2008, pp. 485-491, 2008

[7]. W. Wang, Y. Yu, S. Jiang, Image Retrieval by Emotional Semantics: A Study of Emotional Space and Feature Extraction, *IEEE Systems, Man and Cybernetics (SMC)*, Vol. 4, Issue 8-11, pp. 3534 – 3539, 2006

[8]. R. Plutchik, H. Kellerman, *Emotion: Theory, Research, And Experience*, Academic Press, San Diego, 1990

[9]. R. Kohavi, R. Quinlan, Decision Tree Discovery, *Handbook of Data Mining and Knowledge Discovery*, Klosgen & Zytkow Editors, Oxford University Press, Chapter 16.1.3, pp. 267-276, 2002.

[10]. B. Berlin, P. Kay, *Basic color terms: their universality and evolution*, Berkeley, University of California Press, 1969

[11]. L. Zhang, F. Lin, B. Zhang, A CBIR Method Based on Color-Spatial Feature, In: *Proceedings of the IEEE Region the 10th Conference*, 1999

[12]. WEKA, Machine Learning Platform, <*http://www.cs.waikato.ac.nz/~ml/index.html*>, 2008