Classification of Climbing Grab Movements Measured with Myo Armbands

Peter Schulz, Dario Treffenfeld-Mäder and Jörg Wilhelms

Abstract—**Understanding climbing movements is vital for improving on performing them efficiently. Supporting this process with technology is a relatively new topic. Yet it is seen as having high potential for assisting trainers and experts in their analysis. In this paper, we present a comparison of a set of established classifiers for classifying grab movements typically seen in climbing based on electromyographic signals measured by Myo armbands.**

I. Motivation

Climbing has become a become a popular sport in recent years (*Freizeitklettern ist Mitteldeutschland zunehmend populär*). According to the international federation of sport climbing there are over 25 million active sport climbers (*Key Figures*). As other sports, climbing includes improving motor skills, strength, endurance, knowledge, etc (Kajastila and Hämäläinen 2014). Professional training consist of repetition cycles during which an experienced trainer observes a climber and gives feedback by commenting on movements (Ladha et al. 2013). Video analysis is used occasionally to assist this process. Though, it is rather cumbersome and remains restricted to observation of the visible. Further, research looking into the climbing movements and trying come up with new concerted training approaches is still a rather young discipline, yet (Denk 2015).

Approaches have been made using technology to support the training process, for example full body motion analysis using a Microsoft Kinect (Cha et al. 2015; Kajastila and Hämäläinen 2014). Wearables have also been used for analysing various aspects of climbing movements (Ladha et al. 2013) or have been at least suggested (Kalyanaraman, Ranjan, and Whitehouse 2015). So far no efforts have been made to look into the finger grab movements. The fingers are of special interest since they exhaust quickly and are most vulnerable to injuries (Wright, Royle, and Marshall 2001). Hence analyzing finger movements in sport climbing might help to recognize potentially harmful ones earlier.

II. Concept

A. How the Myo Armbands Work

Electromyography (EMG) is the investigation of the electrical function of the muscles of the human body. If a muscle is contracted there is a measurable muscle response or electrical activity triggered by the nerves of the muscle.

The intramuscular method needs direct access to the nerves (by a needle or other instruments). This method is used by neuromedicine to detect neuromuscular deviation (The Johns Hopkins University 2013). Deviations of the electrical activity can for example indicate diseases of the muscle or nerves. This method is not recommandable for this project, because there is always a risk of infections or other harms to the proband.

On the surface of the muscles you can also measure electrical activity. This method, called surface EMG, is less accurate because it is harder to precisely recognize the muscle that was moved, or the nerve that was used to trigger the muscle activity (Klinik für Klinische Neurophysiologie 2016). The Myo gesture control armband works with surface EMG, but the data delivers appropriate readings for this project.

B. Hypothesis

Hypothesis 1 H_{0_1}: *The classification accuracy is independent of the test subject hence it does not make a difference if only one classifier is used for all test subjects or if one classifier is used per test subject.*

A classifier trained per test subject yields significantly more accurate results than a classifier trained on all test subjects.

Hypothesis 2 H_{0_2}: *The classification accuracy is independent of the pull direction hence it does not make a difference if only one classifier is used for all pull directions or if one classifier is used per pull direction.*

A classifier trained per pull direction yields significantly more accurate results than a classifier trained on all pull directions.

Hypothesis 3 H_{0_3}: *The classification accuracy is independent of the arm hence it does not make a difference if only one classifier is used for both arms or if one classifier is used per arm.*

A classifier trained arm yields significantly more accurate results than a classifier trained on both arms.

C. Getting Data to Test On

1) Record Data: To record data for the evaluation a software must was written for that purpose. This program provides a graphical user interface which allows to manage subjects an there attributes next to recordings. It can visualize the data for visual inspection. It allows for recordings of discrete length.

2) Preprocessing Data: The conducted tests are based on recordings of EMG signals measured and broadcast by a Myo armband at $200\,\mathrm{Hz}$ (Bernhardt 2015). Since Bluetooth LE cannot cope with this frequency, a Myo sends 2 sets of 8 bytes at the maximum rate. Each sample is rather small and could be seen as a 2x8 pixel gray-scale image. As for this we decided to omit a dedicated feature selection and instead

consider each sensor value as a relevant feature for the time being.

However, each sensor value gets preprocessed as follows. 1) all signals are run through an envelope follower filter to get rid of the oscillation and to smooth peaks. 2) The signal stream is split into chunks of n_c samples which are compressed to the mean value of this chunk. 3) Each mean value is mapped to an interval indexed from 0 to n_i. The larger n_c the more samples are considered for one mean value and less instances can be extracted from one recording. When n_i grows the number of possible values for an instance attribute decreases. Choosing appropriate values for n_c and n_i is part of the evaluation process.

3) Data Analysis with Weka[1]: After preprocessing the data it will be given to different weka algorithms to find the best one for climbing gesture recognition. This classifiers are trained on different data sets.

For every test run the chosen data to train the classifiers is different. With the different data sets the best classifier for different subjects, arms and pull directions can be determined.

III. EXPERIMENTAL SETUP

The experiment was executed with 1 person at a time to make sure gestures, tags and arms are assigned correctly. The experiment was executed with 5 probands aged between 18 and 34, 1 proband female, 4 male.

The experiment was split into different test locations. 4 tests were executed as a field test in a climbing hall and 1 test was executed as laboratory test.

At the beginning the proband had to take on 2 Myo gesture control armbands, so the armbands could warm up and the tester could ensure that the assignment of the arm to the armbands was correct. A simple ping in the analysis application was used to determine the correct arm.

The proband had to do all gestures from figure 1 for every pull direction from figure 2. The gestures determine the amount of fingers used (and which fingers are used) and the pull direction determines the direction, in which the proband has to pull.

Before each gesture was started, the tester told the proband what he had to do, let the proband start the current gesture in the current pull direction, started the recording and after the recording was done, told the proband to release the gesture.

IV. IMPLEMENTATION

A. Receiving Data

The Myo SDK (Thalmic Labs Inc. 2016) provides access to several gestures that are recognized. These gestures are not applicable to recognition of climbing processes. The Myo armband also provides raw EMG data from 8 different EMG Sensors with redundant data that allows to evaluate the data by oneself. Each sensor provides a byte value that needs to be processed.

[1]The University of Waikato 2016.

B. Experience

- automatic detection of arm unreliable
- difficult to get reproducible process for all probands

V. EVALUATION

A. Classifiers

A classifier is an algorithm that predicts the type, label, name or class of a given datum. In this project we use algorithms using Naive Bayes, Logistic Regression, Decision Trees and Support Vector Machine to classify the data.(Smola and Vishwanathan 2008)

1) NaiveBayes: The NaiveBayes classifier assigns a data set either the class with the highest probability to be the right on, or the one it's cheapest to assign it to. Computing the likelihood that a dataset is from a given class can be done in linear time and is faster than Logistic Regression, Decision Trees or Support Vector Machine. (Rish 2001)

2) J48 (C4.5 algorithm): J48 is a Java implementation of the C4.5 algorithm. The J48 takes the training data and generates a decision tree from it. The classifier can now classify the test set instance by going down the branches of the decision tree that match the test instance. (Wu 2008)

3) Logistic: The Logistic classifier is the Weka implementation that uses Logistic Regression. Logistic Regression builds a function that calculates the probability for an specific outcome. (Wu 2008)

4) Support Vector Machine: We use LibSVM as our Support Vector Machine classifier. The classes in the training set are divided by an hyperplane $f(x)$. Once the data is divided the test data set x_n is passed to the function, we kann now see the predicted class form the result. (Wu 2008)

B. Data structure

The myo armband sends the values for all of it's 8 emg sensors multiple times a second. For our test we also know the test subject, the pull direction (pull-downwards, pull-upwards, pull-sideways, pull-gaston), the arm (left or right) and the combination of fingers that were used (index, index+middle, index+middle+ring, index+thumb, index+middle+thumb ,index+middle+ring+thumb). We have to translate this data into a data structure supported by Weka. Weka works with `Instances`, `Instance` and `Attribute`. An `Attribute` is in our case either an emg value or a combination of fingers. We combine emg values and the combination of fingers that happed at the same point in time together into an `Instance`. This `Instance` looks something like this:

```
4,3,3,3,3,3,3,4,index
```

The `Instances` contains all of our `Instance` Objects. Weka saves this `Instances` in an *.arff file, which is human readable:

```
@attribute emg_0 numeric
...
@attribute emg_7 numeric
@attribute grip_type {index, ...}
```

(a) index

(b) index+middle

(c) index+middle+ring

(d) index+thumb

(e) index+middle+thumb

(f) index+middle+ring+thumb

Fig. 1: Gestures

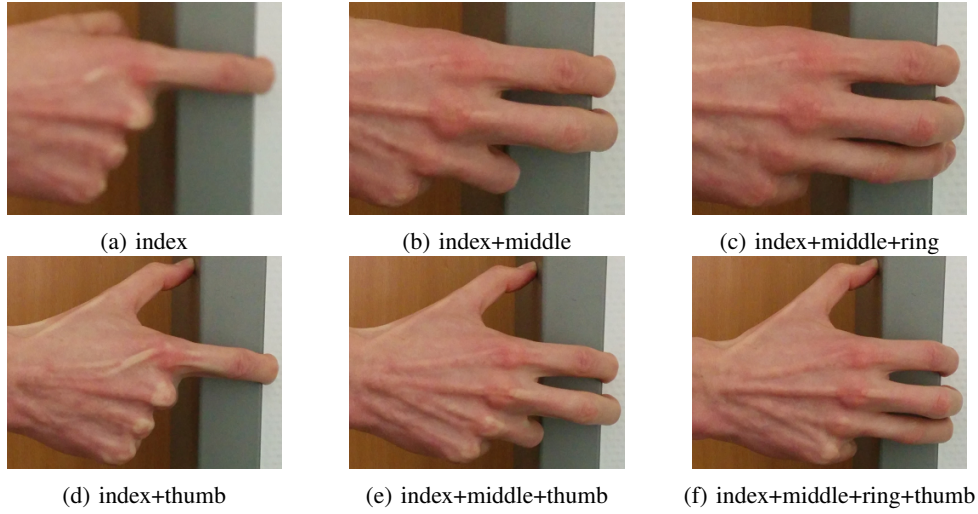(a) pull-downwards

(b) pull-gaston
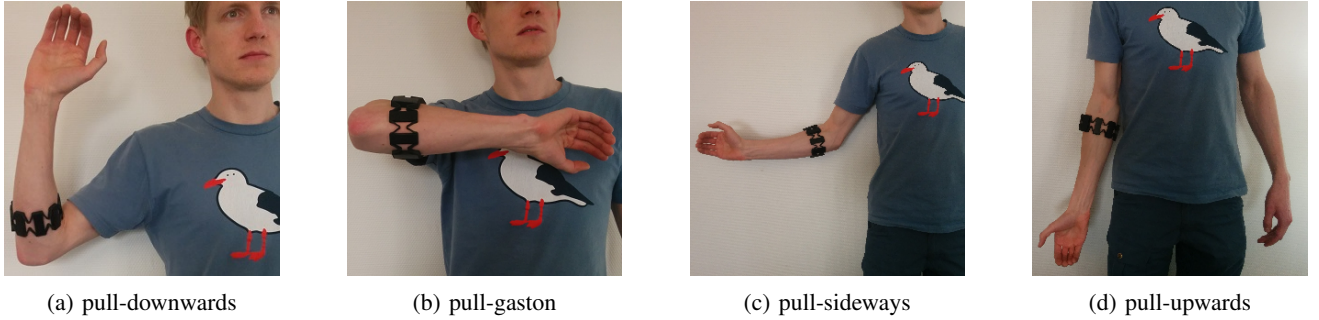
(c) pull-sideways

(d) pull-upwards

Fig. 2: Pull Directions

```
@data
3,3,3,3,3,3,3,3,index
...
4,3,3,3,3,3,3,4,index+middle+ring+thumb
```

This data structure can now be used by Weka.

1) Training set and Test set: Each classifier is trained on a training set and tested against a test set. Both sets are created from the same set of recordings. In order to get two distinct sets of instances each recording is split. The training set is generated of the first 80% of every recording meanwhile the test set comes from the remaining 20%. Other ways Of splitting a recording, for example picking the test portion randomly, have been considered but not tested for the sake of simplicity. The training data gets normalized by subtracting the mean of all values from each value and dividing the difference by the standard deviation of all values. Both mean and standard deviation of the training set are used to normalize the training data, too.

2) Classifier Assessment: In order to compare classifiers they are each assessed separately based on the same data. First, cross Validation is used to measure the accuracy on the training set. Each training set is randomly cross validated using 10 folds. This way an average of 16, at least 8, randomly picked instances is tested against the remaining instances. The test results of all 10 folds are summarized as one classifier evaluation. Second, for a real life scenario each trained classifier is run against the unseen instances of the test set which results in another classifier evaluation.

3) Filtering the Data: Our raw data is filterd by 3 filter before given to the classifier. The first filter is an envelop filter. This filter smooths the data and converts every negative value into a positive one. The second filter is an average filter. It condenses n_c EMG values into one average value. The third filter is a label filter. We give the filter the variable n_i and it converts each EMG value to one of n_i labels.

C. Running the Tests

The tests where run in two phases. During the first phase the aforementioned filter parameters n_c and n_i where determined. These parameters where then used in the second phase where multiple classifiers where trained on different training sets and compared with each other.

1) Filter Parameters: To figure out good chunk size n_c and number of intervals n_i one and the same test was run several times testing different combinations of n_c and n_i. The test was a cross validation of the complete training set including all recordings. We choose n_i to run from 0 (no filtering) to 20 and n_c from 0 (no filtering) to 30. The results vary depending on the classifier, see Table I.

Classifier	n_c	n_i	Accuracy (in %)
LibSVM	15	26	82.65
J48	14	16	80.61
Logistic	15	16	73.46
NaiveBayes	16	16	61.22

TABLE I: Preprocessing filter parameters n_c (chunk size) and n_i (number of intervals)

Test X	Method	Δ Means	t	P
subject	CV	19.808	11.485	<0.001
subject	Test	18.657	8.181	<0.001
pull direction	CV	12.769	7.430	<0.001
pull direction	Test	10.849	4.798	<0.001
arm	CV	5.884	4.460	<0.001
arm	Test	4.720	2.814	0.005

TABLE II: Two-Way ANOVA comparison of cross validation (CV) and test accuracy (in %) based on the factor "number of X in training set". Each row compares one classifiers trained on all X vs on trained per X.

Classifier	Δ Means	t	P
LibSVM	1.752	0.664	0.507
J48	3.850	1.459	0.146
Logistic	4.307	1.632	0.104
NaiveBayes	13.626	5.164	<0.001

TABLE III: Two-Way ANOVA comparison cross validation accuracy (in %) based on number of arms in training set within the examined classifiers
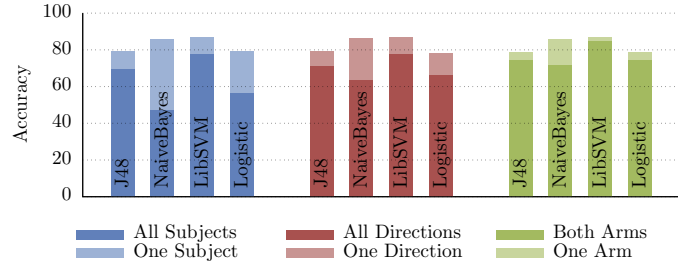

Fig. 3: Two-Way ANOVA comparison of mean test accuracy (in %) for all examined classifiers

2) Hypothesis Testing: In order to test Hypothesis 1 a classifier was trained for each combination of arm and pull direction with recordings from all subjects. Another batch of classifiers was trained for each combination of arm, pull direction, and subject. All evaluation results where analyzed using a Two-Way ANOVA. The number of subjects per training set and the classifier where chosen as factors. This method might be error prone as stated by Salzberg since it ignores that the conducted tests are not independent. Still, the results should suffice as an indicator. Table II shows the results of the comparison. Both, cross validation accuracy and test accuracy off all classifiers examined are significantly higher when the classifier is trained per subject, see also Table V and VI and Figure 3. Regarding the factor "number of subjects per training set", J48 produced the highest p-value with 0.024 followed by LibSVM with 0.011. Whereby a higher p-value indicates a smaller variance between the groups of each factor. This disproves H_{0_1} and thereby supports Hypothesis 1.

Hypothesis 2 was tested in a similar fashion. Notably the difference of means for both cross validation accuracy and test accuracy is almost half a low as for as for the previous test. Which means that there are smaller differences in-between classifiers. Since all examined classifiers produced significantly better results when trained with a more specialized training set Hypothesis 2 becomes unlikely, too.

Finally Hypothesis 3 was tested the same way as the other hypothesis before. At first glance a comparison based on the factor "number of arms per training set" the still shows significant differences between the two possible characteristics. However, looking at the comparison based on the same factor within each classifier, there is almost no significant difference. As can be seen in Table III, only NaiveBayes produces a significantly different cross validation accuracy depending on the number of arms it was trained on. For the test accuracy none of the classifiers produced any significant difference. Hence H_{0_3} might be likely after all.

VI. CONCLUSION AND FUTURE WORK

We have presented a comparison of a set established classifiers for classifying typical grab gestures used in climbing based on EMG signals measured at the underarm. The results imply the chance of classifying a gesture correctly in this scenario are dependent on test subject and pull direction but is invariant with regard to the arm the sensor is worn. Simply put, a set of classifiers trained subject might improve the classification accuracy. If this set contains classifiers which are picked depending on the pull direction, which could be determined based on the IMU signals of the Myo, this might increase the classification accuracy even further. On the other hand, having a classifier for each arm is unlikely to improve the classification accuracy. Looking at Table IV an SVM based classifier appears to produce the most accurate results.

Our findings are just a small step towards a better understanding of movements in climbing and their analysis. We encourage other to delve into this topic, for example, by examining potentially harmful grab movements and try to recognize them. Considering the body weight of subjects would be another option. It could be examined if the weight of a subjects or the portion of that weight lifted while grabbing and pulling has any influence on the classification results. We further propose combining the Myo based grab movement analysis with other sensors for full body motion tracking for obtaining a better understanding of climbing movements, for example, how they influence the position of the center of gravity which plays an important role.

VII. ACKNOWLEDGMENTS

Classifier	Subject CV	Subject Test	Direction CV	Direction Test	Arm CV	Arm Test
LibSVM	82.210	76.468	82.405	78.246	85.999	80.790
J48	74.513	70.016	74.995	70.944	76.816	74.067
Logistic	67.678	64.357	72.171	68.941	76.406	72.367
NaiveBayes	66.751	63.436	74.816	71.763	79.059	74.927
Std Err of LS Mean	1.725	2.281	1.719	2.261	1.319	1.678

TABLE IV: Two-Way ANOVA comparison grouped mean accuracy (in %) of the examined classifiers for different factors. These mean values contain results from H_0 training sets and their counterparts.

| | Subject | | | Direction | | | Arm | | | | |
	All	One	Δ	All	One	Δ	Both	One	Δ	ØΔ	ØΔ σ^2
J48	69,74	79,29	9,56	71,01	78,98	7,97	74,89	78,74	3,85	7,12	2,95
NaiveBayes	47,68	85,82	38,14	63,53	86,10	22,57	72,25	85,87	13,63	24,78	12,41
LibSVM	77,64	86,78	9,14	78,01	86,80	8,78	85,12	86,88	1,75	6,56	4,16
Logistic	56,48	78,88	22,40	66,29	78,05	11,76	74,25	78,56	4,31	12,82	9,09
SEM	3,12	1,47		3,07	1,54		2,12	1,58			

TABLE V: Two-Way ANOVA comparison of mean cross validation accuracy (in %) between the two levels of factor "number of X in test set"

| | Subject | | | Direction | | | Arm | | | | |
	All	One	Δ	All	One	Δ	Both	One	Δ	ØΔ	ØΔ σ^2
J48	64,80	75,23	10,43	66,66	75,30	8,64	72,90	75,23	2,33	7,13	4,26
NaiveBayes	46,49	80,38	33,89	63,14	80,38	17,24	69,47	80,38	10,91	20,68	11,87
LibSVM	70,58	82,36	11,78	74,14	82,36	8,22	79,22	82,36	3,13	7,71	4,34
Logistic	55,09	73,62	18,53	64,26	73,62	9,37	71,11	73,62	2,51	10,14	8,04
SEM	4,13	1,95		4,05	2,02		2,69	2,01			

TABLE VI: Two-Way ANOVA comparison of mean test accuracy (in %) between the two levels of factor "number of X in test set"

REFERENCES

Bernhardt, Paul (2015). URL: http://developerblog.myo.com/myocraft - logging - imu - and - raw - emg - data/ (visited on 11/25/2015).

Cha, Kyungsik et al. (2015). "Analysis of Climbing Postures and Movements in Sport Climbing for Realistic 3D Climbing Animations". In: *Procedia Engineering* 112. 'The Impact of Technology on Sport VI' 7th Asia-Pacific Congress on Sports Technology, {APCST2015}, pp. 52 –57. ISSN: 1877-7058. DOI: http://dx.doi.org/10.1016/j.proeng.2015.07.175. URL: http://www.sciencedirect.com/science/article/pii/S1877705815014241.

Denk, Franziskus (2015). *Gespräch mit Franziskus Denk, Trainer C Sportklettern.*

dpa/sa. *Freizeitklettern ist Mitteldeutschland zunehmend populär.* URL: http : / / www . welt . de / regionales / sachsen - anhalt / article145289159 / Freizeitklettern - ist - Mitteldeutschland - zunehmend - populaer . html (visited on 08/17/2015).

IFSC. *Key Figures.* URL: http://www.ifsc-climbing.org/index.php/media-centre/key-figures-2 (visited on 02/02/2016).

Kajastila, Raine and Perttu Hämäläinen (2014). "Augmented Climbing: Interacting with Projected Graphics on a Climbing Wall". In: *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems.* CHI EA '14. Toronto, Ontario, Canada: ACM, pp. 1279–1284. ISBN: 978-1-4503-2474-8. DOI: 10.1145/2559206.2581139. URL: http://doi.acm.org/10.1145/2559206.2581139.

Kalyanaraman, Avinash, Juhi Ranjan, and Kamin Whitehouse (2015). "Automatic Rock Climbing Route Inference Using Wearables". In: *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers.* UbiComp/ISWC'15 Adjunct. Osaka, Japan: ACM, pp. 41–44. ISBN: 978-1-4503-3575-1. DOI: 10.1145/2800835.2800856. URL: http://doi.acm.org/10.1145/2800835.2800856.

Klinik für Klinische Neurophysiologie (2016). *Elektromyographie und Elektroneurographie.* URL: http://www.neurologie.uni - goettingen . de / index . php / elektromyographie - und - elektroneurographie.html (visited on 02/03/2016).

Ladha, Cassim et al. (2013). "ClimbAX: Skill Assessment for Climbing Enthusiasts". In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* UbiComp '13. Zurich, Switzerland: ACM, pp. 235–244. ISBN: 978-1-4503-1770-2. DOI: 10.1145/2493432.2493492. URL: http://doi.acm.org/10.1145/2493432.2493492.

Rish, Irina (2001). "An empirical study of the naive Bayes classifier". In: *IBM Research Report.*

Salzberg, Steven L. "On Comparing Classifiers: Pitfalls to Avoid and a Recommended Approach". In: *Data Mining and Knowledge Discovery* 1.3, pp. 317–328. ISSN: 1573-756X. DOI: 10.1023/A:1009752403260. URL: http://dx.doi.org/10.1023/A:1009752403260.

Smola, Alex and S.V.N. Vishwanathan (2008). "Introduction to Machine Learning". In: Cambridge: Cambridge University Press.

Thalmic Labs Inc. (2016). *Myo SDK Manual.* URL: https://developer.thalmic.com/docs/api_reference/platform/index.html (visited on 02/03/2016).

The Johns Hopkins University (2013). *Electromyography (EMG)*. URL: http : / / www . hopkinsmedicine . org / healthlibrary / test_procedures / neurological / electromyography_emg_92 , P07656/ (visited on 02/03/2016).

The University of Waikato (2016). *Weka 3: Data Mining Software in Java*. URL: http://www.cs.waikato.ac.nz/ml/weka/ (visited on 02/03/2016).

Wright, D M, T J Royle, and T Marshall (2001). "Indoor rock climbing: who gets injured?" In: *British Journal of Sports Medicine* 35.3, pp. 181–185. DOI: 10.1136/bjsm.35.3.181. eprint: http://bjsm.bmj.com/content/35/3/181.full.pdf+html. URL: http://bjsm.bmj.com/content/35/3/181.abstract.

Wu, Xindong (2008). "Top 10 algorithms in data mining". In: *Knowledge and Information Systems*.