Team: **MLKD**
Final result: **0.527 and 9th place**
Team members:
Eleftherios Spyromitros-Xioufis (contact at: espyromi@csd.auth.gr)
Grigorios Tsoumakas
Ioannis Vlahavas
Department of Informatics, Aristotle University of Thessaloniki, Greece

Brief Description of the Approach

We first applied feature selection according to the x^2_{max} criterion [1] in order to reduce the size of the feature space and to select the most informative feature variables. This process led to training and test datasets of variable feature space sizes (1000, 2000,…,23000). Next, we followed the *Binary Relevance* approach for multi-label classification where a separate binary model is learned for each label using traditional binary classifiers. We evaluated different types of classifiers by applying 3-fold cross-validation on the training set and measuring the average F-measure of the predictions. In all cases, we applied an approximate F-measure optimization method based on [2] to transform the probability outputs of the classifiers into hard 0/1 classifications. After experimentation we found that the best performing base classifiers were *Random Forest* and *L2-regularized Logistic Regression*. The final solution comes by averaging the probability outputs of the two methods and then applying the F-measure optimization method. However, we would like to mention that the performance gain of the combination of the two classifiers was minimal and that L2-regularized Logistic Regression was at least an order of magnitude faster than Random Forest during training.

The Random Forest model was learned on the full training set (no feature selection applied) using 200 trees. The L2-regularized logistic regression model was learned on the top 13000 features. The evaluation was carried out using the multi-label learning library Mulan[1] which provides implementations of the Binary Relevance and the x^2_{max} feature selection methods. We modified the library for F-measure optimization and calculation of F-measure according to the contest's definition. Finally we used Weka's[2] implementation of Random Forest and LIBLINEAR's[3] implementation of L2-regularized Logistic Regression.

[1] David D. Lewis, Yiming Yang, Tony G. Rose, and Fan Li. 2004. RCV1: A New Benchmark Collection for Text Categorization Research. *J. Mach. Learn. Res.* 5 (December 2004), 361-397.
[2] David D. Lewis. 1995. Evaluating and optimizing autonomous text classification systems. In *Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval* (SIGIR '95), Edward A. Fox, Peter Ingwersen, and Raya Fidel (Eds.)

[1] http://mlkd.csd.auth.gr/multilabel.html
[2] http://www.cs.waikato.ac.nz/ml/weka/
[3] http://www.csie.ntu.edu.tw/~cjlin/liblinear/