

Suppliment 2 to CLRC—TR—08—01

UKOPS: Supplementary Results 2

Dmitry Devetyarov, Iliia Nouretdinov, Brian Burford, Zhiyuan Luo,
Alexey Chervonenkis, Volodya Vovk, Alex Gammerman
Royal Holloway, University of London

January 17, 2009

Abstract

The supplement provides the results of the analysis of UKOPS Reading data set randomly divided into new training and test sets.

The division of UKOPS data into training and test sets analysed in previous research was biased, since in the test set, all Malignant samples had CA125 level higher than 30 and only one Healthy sample has CA125 higher than 30. Thus, Healthy and Malignant classes were separable by CA125 cutoff, while we could observe a big overlap in CA125 distributions for Healthy and Malignant classes in the training set .

For this reason, UKOPS training and test sets were merged together and randomly divided into new training and test sets. They were divided so that the ratio of each class (Healthy, Benign, Malignant, Borderlines) in the training set and test set is 2 : 1.

The Tables 1 and 2 demonstrate the results for Healthy vs Malignant for both $Quality = (2 * Sensitivity + Specificity)/3$ and $Quality = (Sensitivity + 2 * Specificity)/3$. In brief, among models that outperform CA cut-off model on the training set, there are models that outperform this model on test set (by 1 correctly classified healthy sample), and there are models that match the cut-off model. The rules for different Qualities (weighting towards Sensitivity and Specificity) are almost the same.

Table 1: Results for random division into training and test sets, where Quality = (Sensitivity + 2 * Specificity)/3

Rule	Peak 1, Da	Peak 2, Da	Training Set				Test Set			
			Sens	Spec	Acc	Qual	Sens	Spec	Acc	Qual
CA125>30	—	—	97.7%	97.3%	97.5%	97.5%	95.7%	96.5%	96.3%	96.2%
1 out of 2	1466.9	5487	100.0%	99.1%	99.4%	99.4%	95.7%	94.7%	95.0%	95.0%
1 out of 3	2660.9	8438.6	100.0%	99.1%	99.4%	99.4%	95.7%	94.7%	95.0%	95.0%
2 out of 7	852.93	8438.6	100.0%	99.1%	99.4%	99.4%	95.7%	96.5%	96.3%	96.2%
2 out of 8	852.93	8309.3	100.0%	99.1%	99.4%	99.4%	95.7%	93.0%	93.8%	93.9%
1 out of 1	2755.2	1897.7	97.7%	100.0%	99.4%	99.2%	95.7%	96.5%	96.3%	96.2%
2 out of 4	2660.9	808.26	97.7%	100.0%	99.4%	99.2%	95.7%	98.3%	97.5%	97.4%
2 out of 5	2660.9	852.93	97.7%	100.0%	99.4%	99.2%	95.7%	96.5%	96.3%	96.2%
2 out of 6	2755.2	805.13	97.7%	100.0%	99.4%	99.2%	95.7%	98.3%	97.5%	97.4%
3 out of 10	2660.9	852.93	97.7%	100.0%	99.4%	99.2%	95.7%	98.3%	97.5%	97.4%
3 out of 7	2660.9	808.26	97.7%	100.0%	99.4%	99.2%	95.7%	98.3%	97.5%	97.4%
3 out of 8	7219.2	8565.3	97.7%	100.0%	99.4%	99.2%	95.7%	98.3%	97.5%	97.4%
3 out of 9	2660.9	852.93	97.7%	100.0%	99.4%	99.2%	95.7%	98.3%	97.5%	97.4%
4 out of 10	2660.9	808.26	97.7%	100.0%	99.4%	99.2%	95.7%	98.3%	97.5%	97.4%

Table 2: Results for random division into training and test sets, where Quality = (2 * Sensitivity + Specificity)/3

Rule	Peak 1, Da	Peak 2, Da	Training Set				Test Set			
			Sens	Spec	Acc	Qual	Sens	Spec	Acc	Qual
CA125>30	—	—	97.7%	97.3%	97.5%	97.6%	95.7%	96.5%	96.3%	95.9%
2 out of 7	852.9	8438.6	100.0%	99.1%	99.4%	99.7%	95.7%	96.5%	96.3%	95.9%
1 out of 2	1466.9	5487	100.0%	99.1%	99.4%	99.7%	95.7%	94.7%	95.0%	95.4%
1 out of 3	2660.9	8438.6	100.0%	99.1%	99.4%	99.7%	95.7%	94.7%	95.0%	95.4%
2 out of 8	852.9	8309.3	100.0%	99.1%	99.4%	99.7%	95.7%	93.0%	93.8%	94.8%
2 out of 4	2660.9	808.26	97.7%	100.0%	99.4%	98.5%	95.7%	98.3%	97.5%	96.5%
3 out of 10	2660.9	852.93	97.7%	100.0%	99.4%	98.5%	95.7%	98.3%	97.5%	96.5%
3 out of 7	2660.9	808.26	97.7%	100.0%	99.4%	98.5%	95.7%	98.3%	97.5%	96.5%
3 out of 8	7219.2	8565.3	97.7%	100.0%	99.4%	98.5%	95.7%	98.3%	97.5%	96.5%
3 out of 9	2660.9	852.9	97.7%	100.0%	99.4%	98.5%	95.7%	98.3%	97.5%	96.5%
4 out of 10	2660.9	808.3	97.7%	100.0%	99.4%	98.5%	95.7%	98.3%	97.5%	96.5%
1 out of 1	2755.2	1897.7	97.7%	100.0%	99.4%	98.5%	95.7%	96.5%	96.3%	95.9%
2 out of 5	2660.9	852.9	97.7%	100.0%	99.4%	98.5%	95.7%	96.5%	96.3%	95.9%