Multi-Class Probabilistic Active Learning

Daniel Kottke and **Georg Krempl** and
Dominik Lang and **Johannes Teschner** and **Myra Spiliopoulou**

Abstract

This work addresses active learning for multi-class classification. Active learning algorithms optimize classifier training by subsequently selecting those instances for labeling by an expert, which improve the classifier's performance the most. In this work, we identify different influence factors that positively affect active learning. These factors are (1) an instance's impact, (2) its posterior, and (3) the reliability of this posterior. We contribute a new decision-theoretic approach, called multi-class probabilistic active learning (McPAL). Building on a probabilistic active learning framework, our approach is non-myopic, fast, and optimizes a performance measure (like accuracy) directly. Considering all influence factors, McPAL determines the expected gain in performance to compare the usefulness of instances. For this purpose, it calculates the density weighted expectation over the true posterior and over all possible labeling combinations in a closed-form solution. Thus, in contrast to other multi-class algorithms, it considers the posterior's reliability which improved the performance. In our experimental evaluation, we show the reasonability of the selected influence factors and the superiority of McPAL in comparison to various other multi-class active learning algorithms on six datasets.

SUPPLEMENTAL MATERIAL

Table 1. Mean error and standard deviation of the all algorithms on our six datasets using the Parzen window classifier. We report the results after 20, 40, and 60 acquired labels. The best method is printed in bold numbers.

20 samples	ecoli	glass	iris	vehicle	wine	yeast
McPAL	**22.70** (± *4.45*)	**30.17** (± *4.22*)	**3.94** (± *1.97*)	149.14 (± *11.94*)	**2.66** (± *1.43*)	**275.24** (± *26.35*)
BvsSB	24.75 (± *4.84*)	35.95 (± *5.57*)	12.63 (± *7.06*)	148.68 (± *18.25*)	2.80 (± *1.67*)	289.90 (± *23.13*)
MaxECost	25.42 (± *6.63*)	33.33 (± *5.09*)	8.23 (± *6.24*)	155.98 (± *17.71*)	2.95 (± *1.88*)	294.20 (± *32.95*)
Conf	24.64 (± *7.07*)	33.93 (± *5.02*)	12.48 (± *7.32*)	156.52 (± *17.19*)	2.90 (± *1.79*)	292.92 (± *34.42*)
Entr	26.94 (± *8.01*)	33.04 (± *5.50*)	14.61 (± *3.17*)	153.44 (± *18.82*)	3.41 (± *1.76*)	298.60 (± *32.63*)
VoI	40.14 (± *9.59*)	38.20 (± *3.98*)	16.55 (± *2.67*)	NA	2.89 (± *2.68*)	NA
Rand	32.52 (± *7.89*)	36.69 (± *5.00*)	9.91 (± *4.47*)	**145.38** (± *13.27*)	4.35 (± *3.04*)	300.12 (± *23.56*)
40 samples	ecoli	glass	iris	vehicle	wine	yeast
McPAL	**19.15** (± *4.06*)	**29.14** (± *4.22*)	**2.85** (± *1.58*)	125.88 (± *8.99*)	1.78 (± *1.06*)	**258.36** (± *24.40*)
BvsSB	21.02 (± *4.42*)	32.28 (± *4.36*)	11.78 (± *7.78*)	**122.90** (± *14.43*)	1.92 (± *1.26*)	273.52 (± *22.95*)
MaxECost	20.80 (± *4.10*)	29.70 (± *4.46*)	7.70 (± *6.44*)	131.82 (± *14.44*)	1.90 (± *1.16*)	274.54 (± *30.65*)
Conf	19.60 (± *4.30*)	29.79 (± *4.87*)	11.69 (± *7.79*)	133.56 (± *14.90*)	1.94 (± *1.19*)	276.36 (± *32.40*)
Entr	23.55 (± *4.80*)	30.64 (± *4.61*)	13.88 (± *3.49*)	139.02 (± *18.57*)	**1.77** (± *1.14*)	284.38 (± *28.05*)
VoI	41.46 (± *7.22*)	38.06 (± *3.78*)	16.74 (± *2.58*)	NA	1.92 (± *1.89*)	NA
Rand	29.80 (± *6.57*)	34.57 (± *5.18*)	8.28 (± *4.03*)	129.88 (± *13.31*)	2.65 (± *1.61*)	281.84 (± *25.48*)
60 samples	ecoli	glass	iris	vehicle	wine	yeast
McPAL	**18.41** (± *3.69*)	**27.08** (± *3.95*)	**5.81** (± *2.54*)	115.26 (± *7.60*)	1.63 (± *1.06*)	**244.12** (± *20.71*)
BvsSB	19.69 (± *4.44*)	29.71 (± *4.22*)	12.71 (± *7.64*)	**113.42** (± *9.95*)	1.76 (± *1.13*)	259.68 (± *22.66*)
MaxECost	20.29 (± *4.55*)	27.99 (± *4.25*)	8.12 (± *5.62*)	120.06 (± *12.42*)	1.66 (± *1.03*)	257.60 (± *26.75*)
Conf	19.91 (± *4.29*)	28.46 (± *4.59*)	12.40 (± *7.59*)	122.34 (± *13.39*)	1.62 (± *1.12*)	259.98 (± *25.76*)
Entr	22.54 (± *4.55*)	31.65 (± *4.91*)	11.94 (± *4.07*)	126.06 (± *14.60*)	**1.53** (± *1.00*)	272.44 (± *24.93*)
VoI	34.20 (± *5.78*)	37.22 (± *4.72*)	15.06 (± *3.49*)	NA	1.54 (± *1.22*)	NA
Rand	28.32 (± *5.65*)	33.55 (± *5.17*)	6.92 (± *2.76*)	123.28 (± *13.26*)	2.30 (± *1.43*)	276.42 (± *26.98*)

Table 2. Mean error and standard deviation of the all algorithms on our six datasets using the pKNN classifier. We report the results after 20, 40, and 60 acquired labels.

Algorithm	ecoli	glass	iris	vehicle	wine	yeast
McPAL	21.08 (± *4.23*)	30.92 (± *5.41*)	4.24 (± *2.22*)	124.56 (± *9.74*)	2.88 (± *1.51*)	267.42 (± *15.91*)
BvsSB	24.31 (± *4.99*)	35.61 (± *5.42*)	12.20 (± *7.62*)	152.06 (± *19.46*)	2.95 (± *1.73*)	289.02 (± *22.45*)
MaxECost	24.91 (± *7.43*)	33.02 (± *5.15*)	8.35 (± *6.02*)	155.46 (± *17.68*)	3.01 (± *1.79*)	301.96 (± *31.20*)
Conf	22.82 (± *6.16*)	34.25 (± *6.06*)	12.48 (± *7.38*)	155.72 (± *18.01*)	2.93 (± *1.89*)	298.76 (± *34.90*)
Entr	25.45 (± *8.02*)	32.61 (± *5.66*)	9.56 (± *5.42*)	155.82 (± *17.95*)	3.49 (± *1.73*)	300.38 (± *29.50*)
VoI	34.94 (± *8.96*)	37.90 (± *4.01*)	15.43 (± *3.39*)	NA	2.99 (± *2.88*)	NA
Rand	30.75 (± *6.71*)	36.38 (± *6.15*)	9.96 (± *4.76*)	147.52 (± *15.32*)	4.80 (± *3.32*)	291.02 (± *25.18*)
60 samples	ecoli	glass	iris	vehicle	wine	yeast
McPAL	18.96 (± *3.74*)	28.13 (± *5.21*)	2.58 (± *1.31*)	110.04 (± *9.17*)	1.91 (± *1.11*)	251.42 (± *14.95*)
BvsSB	20.12 (± *3.44*)	29.71 (± *5.00*)	6.55 (± *7.30*)	127.76 (± *18.35*)	2.28 (± *1.27*)	262.22 (± *21.81*)
MaxECost	19.28 (± *4.21*)	28.90 (± *5.13*)	3.16 (± *2.60*)	130.42 (± *12.93*)	2.12 (± *1.27*)	266.22 (± *24.65*)
Conf	18.72 (± *4.07*)	29.66 (± *5.02*)	6.63 (± *7.28*)	130.92 (± *13.23*)	2.21 (± *1.22*)	268.22 (± *25.92*)
Entr	19.46 (± *3.87*)	30.64 (± *4.48*)	3.27 (± *2.57*)	136.22 (± *18.16*)	2.14 (± *1.17*)	277.86 (± *27.92*)
VoI	24.70 (± *5.38*)	34.82 (± *4.76*)	5.97 (± *2.77*)	NA	1.71 (± *1.45*)	NA
Rand	24.73 (± *5.23*)	31.77 (± *5.14*)	5.19 (± *2.84*)	128.92 (± *14.71*)	2.85 (± *1.62*)	267.84 (± *22.42*)
60 samples	ecoli	glass	iris	vehicle	wine	yeast
McPAL	17.77 (± *3.23*)	26.19 (± *3.96*)	2.65 (± *1.40*)	103.92 (± *9.02*)	1.83 (± *1.21*)	245.74 (± *13.86*)
BvsSB	18.74 (± *3.22*)	26.76 (± *4.65*)	2.59 (± *1.40*)	113.04 (± *8.86*)	2.27 (± *1.23*)	248.36 (± *17.93*)
MaxECost	18.73 (± *4.09*)	25.86 (± *4.63*)	2.47 (± *1.49*)	118.06 (± *9.83*)	2.31 (± *1.25*)	256.08 (± *18.84*)
Conf	18.40 (± *3.94*)	26.68 (± *4.14*)	2.58 (± *1.39*)	118.88 (± *10.35*)	2.28 (± *1.27*)	255.76 (± *19.12*)
Entr	17.69 (± *3.90*)	27.33 (± *4.40*)	2.46 (± *1.44*)	121.08 (± *12.82*)	2.28 (± *1.30*)	261.48 (± *21.50*)
VoI	20.02 (± *4.41*)	32.37 (± *4.72*)	3.07 (± *1.77*)	NA	1.61 (± *1.28*)	NA
Rand	21.93 (± *4.26*)	28.12 (± *4.61*)	3.52 (± *2.12*)	119.90 (± *13.52*)	2.70 (± *1.47*)	254.50 (± *22.11*)

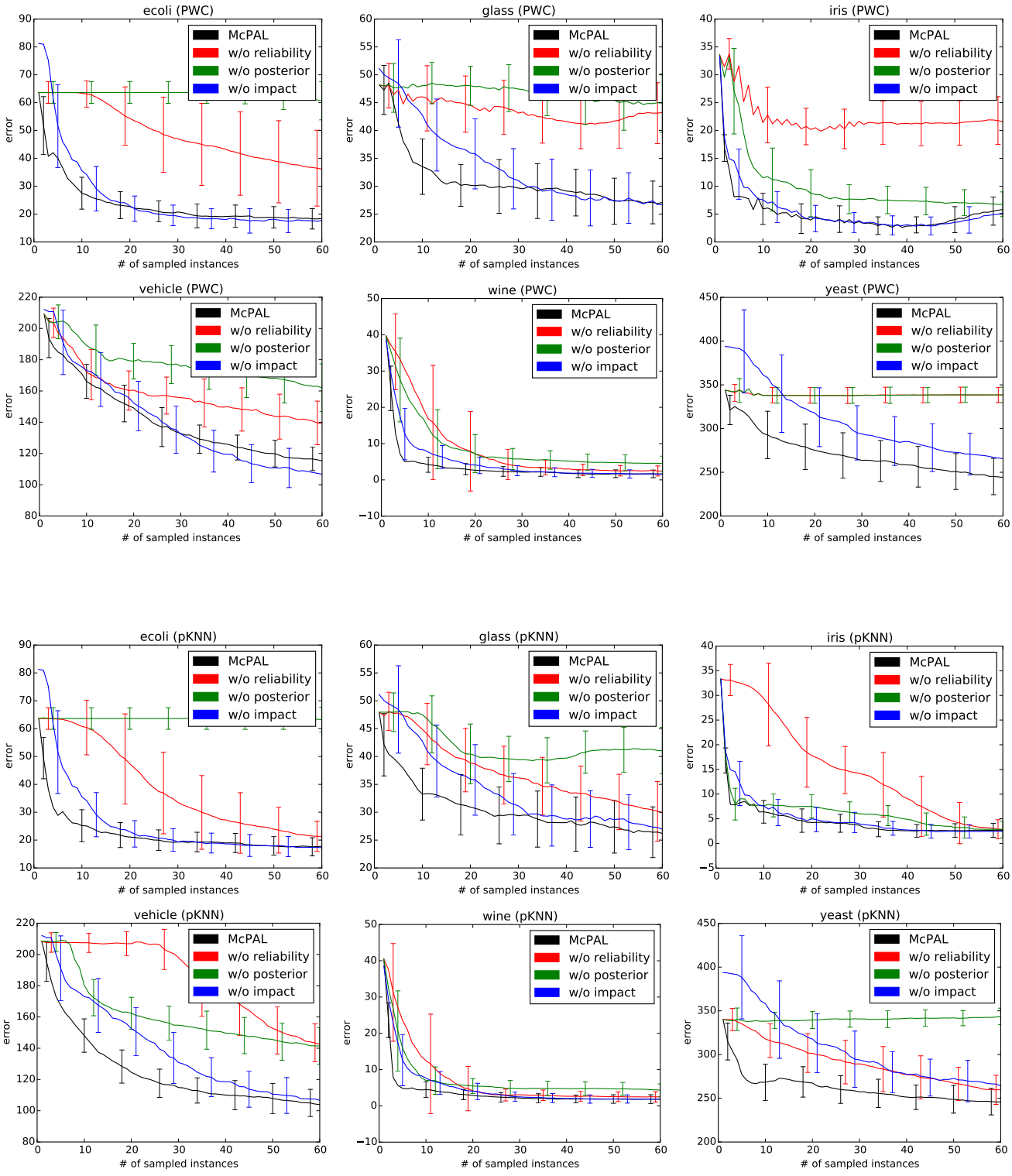

Figure 1. Learning curves of different variants of the McPAL algorithm on all six datasets. The upper plots show results from the PWC classifier, the lower ones with the pKNN.

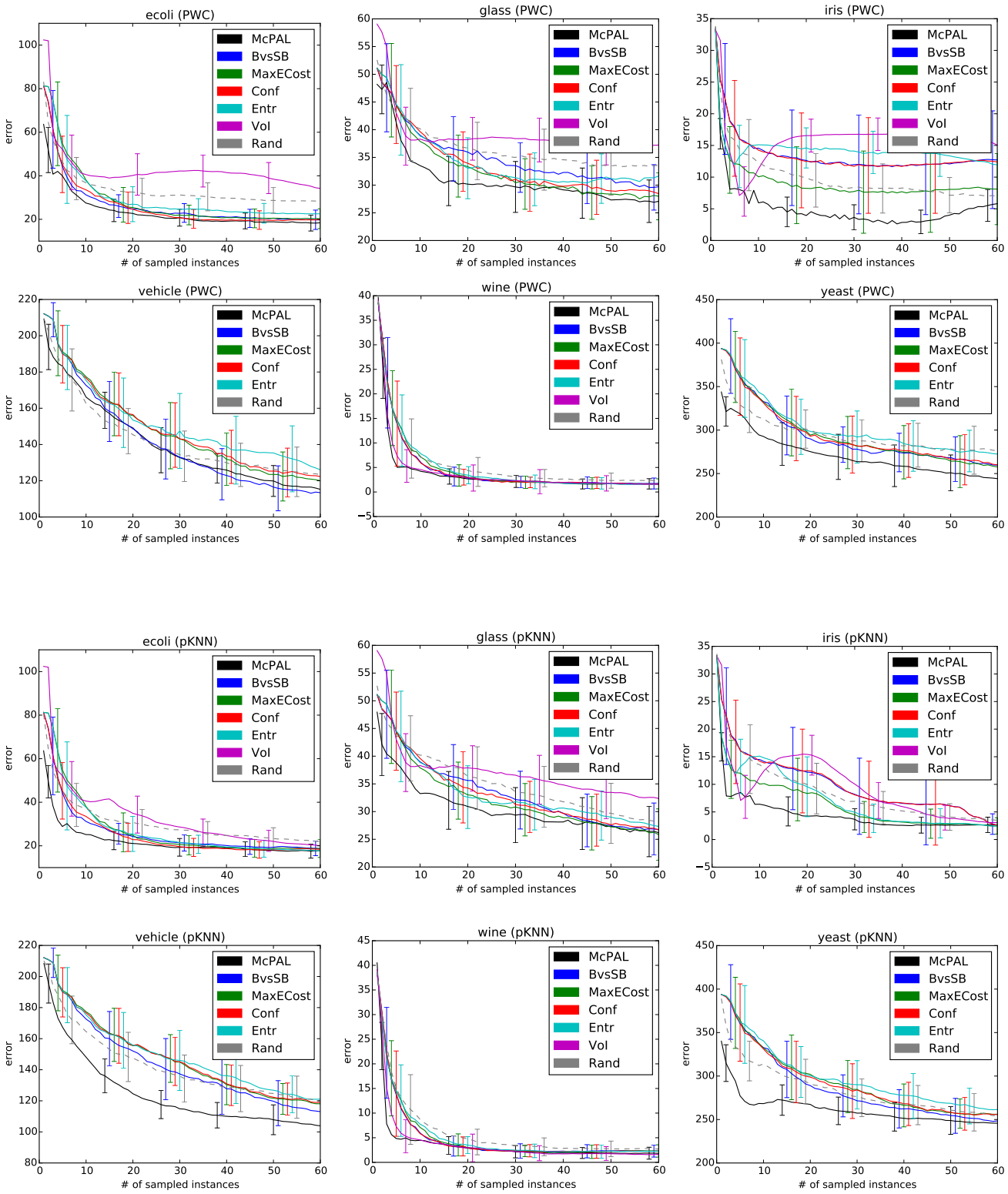

Figure 2. Learning curves of McPAL and its competitors on all six datasets using the PWC and the pKNN classifier.