

Comparison between Different Classifiers in AFE

HE JIANG, Dalian University of Technology

JINGXUAN ZHANG, Nanjing University of Aeronautics and Astronautics

XIAOCHEN LI, Dalian University of Technology

ZHILEI REN, Dalian University of Technology

DAVID LO, Singapore Management University

XINDONG WU, University of Vermont

ZHONGXUAN LUO, Dalian University of Technology

Motivation. We propose AFE to extract features from app descriptions to assist the task of recommending features from mobile app descriptions. AFE consists of three components, namely data cleaner, linguistic rule filter, and feature classifier. In the feature classifier component, we use Naïve Bayes (NB) as the default classifier in the paper and do not compare it with other commonly used classifiers. This technical report compares NB with other commonly used classifiers to show the effectiveness of NB.

Approach. We select other four commonly used classifiers, which are also implemented in Weka, namely Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), and AdaBoost. Among them, DT and SVM are single classifiers. While, RF and AdaBoost are ensemble learning classifiers. By comparing NB with these commonly used classifiers, we can know whether our choice is reasonable.

The same as the paper, we compare different classifiers in the selected five categories. We show some performance indicators for all the categories, such as *Precision*, *Recall*, *F-Measure*, and *Accuracy*.

Results. Table I - IV show the *Precisions*, *Recalls*, *F-Measures*, and *Accuracies* of different classifiers. From Table I we can see that, NB achieves the second-best *Precision* (61.79%) among the commonly used classifiers, which is only less than RF by 1.43%. The *Precisions* of other classifiers are all less than NB. We can find similar phenomenon in the *Recalls* of different classifiers. When comparing *F-Measures* of different classifiers, we can see that NB achieves the best results. For example, NB achieves *F-Measure* of 69.11%, which is the best. Hence, NB can balance *Precision* and *Recall* to achieve better results. From the perspective of *Accuracy*, NB still achieves the best results, i.e., 69.57%.

Conclusion. NB achieves the best results among the commonly used classifiers. Selecting NB as the default classifier in AFE is reasonable.

Table I. Precisions of different classifiers

Category	NB	DT	SVM	RF	AdaBoost
Business	76.19%	69.23%	77.27%	76.60%	69.57%
Education	66.88%	65.00%	66.00%	70.91%	63.30%
Health and Fitness	56.61%	51.61%	55.45%	55.43%	52.80%
Finance	66.19%	62.07%	67.80%	69.16%	63.98%
Music and Audio	43.08%	42.06%	41.12%	44.00%	38.82%
Average	61.79%	58.00%	61.53%	63.22%	57.69%

Table II. Recalls of different classifiers

Category	NB	DT	SVM	RF	AdaBoost
Business	82.05%	78.26%	73.91%	62.61%	83.48%
Education	82.03%	81.25%	77.34%	60.94%	92.97%
Health and Fitness	77.00%	80.00%	61.00%	51.00%	85.00%
Finance	75.41%	73.77%	65.57%	60.66%	84.43%
Music and Audio	84.85%	80.30%	66.67%	50.00%	89.39%
Average	80.27%	78.72%	68.90%	57.04%	87.05%

Table III. F-Measures of different classifiers

Category	NB	DT	SVM	RF	AdaBoost
Business	79.01%	73.47%	75.56%	68.90%	75.89%
Education	73.67%	72.22%	71.22%	65.55%	75.32%
Health and Fitness	65.25%	62.75%	58.10%	53.13%	65.13%
Finance	70.50%	67.42%	66.67%	64.63%	72.79%
Music and Audio	57.14%	55.21%	50.87%	46.81%	54.13%
Average	69.11%	66.21%	64.48%	59.80%	68.65%

Table IV. Accuracies of different classifiers

Category	NB	DT	SVM	RF	AdaBoost
Business	75.94%	69.34%	74.06%	69.34%	71.23%
Education	73.40%	71.63%	71.63%	70.92%	72.34%
Health and Fitness	67.46%	62.30%	65.08%	64.29%	63.89%
Finance	69.92%	66.02%	68.75%	68.36%	69.92%
Music and Audio	61.11%	60.19%	60.65%	65.28%	53.70%
Average	69.57%	65.89%	68.03%	67.64%	66.22%