

CHAPTER 1

Results

1.1 Classifier Experiment Set 1

Experimental results from the first set of experiments were overall excellent, results from them can be seen in Table 1.1 on the following page. In the unbalanced results, an accuracy of 82.9% was observed, and in the balanced results, 78.5%. In the top balanced classifier, J48, the Root Mean Squared Error (RMSE) was 0.28 persons, a result better than Thermosense’s 0.35 persons.

Between the unbalanced and balanced classes, the ranking of different algorithms remained approximately the same, and consistently dropped in accuracy, with the exceptions being SMO, which increased in accuracy by 0.5% in that instance, and, more notably, IBk, which increased in accuracy by 11.2%. The drop in accuracy can be explained mostly by an over-representation of the zero class within the under-balanced data, and an underrepresentation of the three class (see Figure 1.1 on page 3). These biases would enable classes to over-predict and under-predict these two classes respectively and achieve an artificially higher accuracy as a result. As discussed in the Methods, we performed re-sampling inside Weka to compensate for this.

For the numeric representation of the number of people, accuracy was consistently poor. In this instance, we calculate the “correctly classified” percentage by taking the predicted vs. actual results from Weka and rounding the predicted values to whole numbers, as a theoretical real-world solution may do to finalize its prediction. From this data, we can see that all three classifiers used performed consistently poorly, with the Root Mean Square Errors being consistently double or more of comparable nominal results. Interestingly, IBk was on-par with its nominal unbalanced partner, while the Multi-Layer Perceptron’s accuracy dropped by more than 30%.

For the Linear Regression model, the most striking difference to the Thermosense paper is evident. While Thermosense claims an RMSE of 0.409, our results show a much poorer 0.659, a difference of more than 46%. Looking more

Classifier	% Correct	RMS Error	F-Measure
<i>Nominal Balanced</i>			
<i>J48</i>	82.9	0.2878	0.824
<i>KStar</i>	82.6	0.2853	0.818
<i>Multilayer Perceptron</i>	78.5	0.2936	0.777
<i>Naive Bayes</i>	66.2	0.3516	0.644
<i>IBk</i>	57.6	0.4245	0.563
<i>SMO</i>	57.5	0.3795	0.554
<i>Nominal Unbalanced</i>			
<i>J48</i>	78.5	0.2771	0.784
<i>KStar</i>	74.4	0.2926	0.743
<i>MLP</i>	71.9	0.3285	0.718
<i>IBk</i>	68.8	0.3715	0.670
<i>Naive Bayes</i>	61.1	0.3576	0.605
<i>SMO</i>	58.0	0.3853	0.569
<i>Numeric</i>			
<i>Linear Regression</i>	63.4	0.6589	-
<i>IBk</i>	55.8	1.1947	-
<i>Multilayer Perceptron</i>	50.2	0.7768	-

Table 1.1: Classifier Experiment Set 1 Results

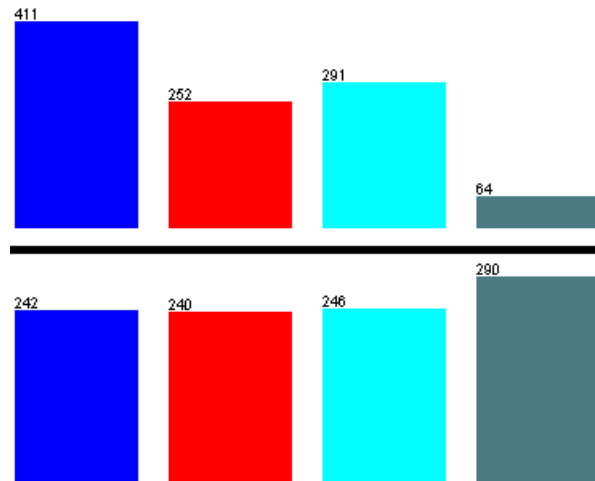


Figure 1.1: Experiment Set 1 Class Distribution Before and After Re-sampling

closely at the Linear Regression generated, Equation (1.1) on page 3, we can see that the impact of the three features on the model are minimal, with the algorithm choosing minuscule weights. This suggests that the algorithm was unable to find any strong correlation between the three features and the number of people.

$$n = 0.0783a + -0.0616b + -0.0331c + 0.4923 \quad (1.1)$$

Bibliography