Project 2: Classifying and Predicting Human Activity Recognition Data

Rahrooh, Allen
Department of Statistics & Data Science
University of Central Florida
Orlando, United States of America
rahrooan@knights.ucf.edu

***Abstract*— Human Recognition Data has been used in the past to classify and predict activity of a human subject. The purpose of this study is to see if a reduction in predictors will have an impact on the classification performance metrics. I used the classification methods of Naïve Bayes, Multinomial Logistic Regression, and Support Vector Machines with a Linear Kernel. I found that with the reduction in predictors that the classification performance decreased by 35% for precision and recall. For the false positive rate I had a 6000% increase which shows that with a reduction in predictors that the classification performance decreases.**

I. Introduction & Background

Classification methods such as Naïve Bayes, Multinomial Logistic Regression, and Support Vector Machines have been used on time series data to predict walking speeds from electroencephalographs, which are electrocortical brain signals. [2].

The paper [1] was interested in the classification and predicating Human Activity Recognition (HAR) data. The HAR dataset included the predictors (X) of the subject's age in years, height in meters, weight in kilograms, body mass index (BMI), and the x, y, z components of accelerometer data mounted on the waist, left thigh, right ankle, and right upper-arm. The dataset also included five different classes which were sitting, sitting down, standing, standing up, and walking. These classes are used as the response variable (Y). After conducting the classification and prediction the performance metrics of the True Positive Rate, False Positive Rate, Precision, Recall, F-Measure, and Receiver Operator Curve (ROC) Area were calculated [1].

II. Methods

To conduct a classification analysis of the HAR data I decided to use R is used to read in the data. For my classification analysis I only focused on the x, y, and z components of the four accelerometers to see if I could get the same results as [1] did because the subject's age, height, weight, and BMI should not reduce the classification performance. Once the data was cleaned and sorted by class I split the data into two datasets which are 80% training and 20% testing. The training dataset is used for the classification analysis and the testing dataset is used for the prediction. I used the popular classification models of Naïve Bayes, Multinomial Logistic Regression, and Support Vector Machines with a Linear Kernel. Once the classification and prediction analysis is performed the results of the classification and prediction is shown using a confusion matrix. A confusion matrix shows how well the model classified the dataset based on the given classes. From the confusion matrix I calculated the True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) for each class and then took an average to get an overall model performance. The TP represents a successful classification of the positive class. The FP represents a failed classification of the positive class. The TN represents a successful negative classification. The FN represents a failed classification of the negative class.

Actual	Predicted	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Fig. 1 Visual Representation of TP, FP, TN, FN

From these values I calculated the accuracy, precision, sensitivity, F1 Score, specificity, negative predicted value (NPV), false negative rate (FNR), and the false positive rate (FPR). Equations (1-8) show the formulas per each metric.

$$\text{Accuracy} = \frac{\text{TP+FN}}{\text{TP+FP+FN+TN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{2}$$

$$\text{Recall (Sensitivity)} = \frac{\text{TP}}{\text{TP+FN}} \tag{3}$$

$$\text{F1 Score} = 2 * \frac{\text{Precision x Sensitivity}}{\text{Precision+Sensitivity}} \tag{4}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \tag{5}$$

$$\text{NPV} = \frac{\text{TN}}{\text{TN+FN}} \tag{6}$$

$$\text{FNR} = \frac{\text{FN}}{\text{FN+TP}} \tag{7}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP+TN}} \tag{8}$$

The accuracy measures how accurate the model is at predicting the actual classes. The precision or positive predicted value (PPV) measures the positive classification performance. The sensitivity measures how well the positive class are correctly identified. The F1 score represents the weighted average of precision and sensitivity. The specificity measures how well the negative class are correctly identified. The NPV measures the negative classification performance. The FNR measures how well the model classifies FN. The FPR measures how well the model classifies FP. For accuracy, precision, sensitivity, specificity, F1 score, and NPV the closer the value is to 1 the better the model is. For FNR and FPR the closer the value is to 0 the better the model is. I then compared the results to the calculated results that [1] reported.

III. Results

Naïve Bayes

Actual	Predicted				
	Sitting	Sitting Down	Standing	Standing Up	Walking
Sitting	9184	100	723	72	5
Sitting Down	222	1341	655	35	116
Standing	59	120	8752	34	510
Standing Up	214	517	1041	329	334
Walking	148	323	1864	519	5910

Multinomial Logistic Regression

Actual	Predicted				
	Sitting	Sitting Down	Standing	Standing Up	Walking
Sitting	10040	19	2	23	0
Sitting Down	137	1375	516	285	56
Standing	0	51	8404	13	1007
Standing Up	107	302	401	1449	176
Walking	5	160	2236	193	6168

Support Vector Machine

Actual	Predicted				
	Sitting	Sitting Down	Standing	Standing Up	Walking
Sitting	10017	17	0	50	0
Sitting Down	152	1459	432	226	100
Standing	0	89	8719	11	656
Standing Up	81	269	397	1452	236
Walking	2	151	2015	96	6500

Fig. 2 Confusion Matrices of the Classifiers

Looking at the confusion matrices for the Naïve Bayes, Multinomial Logistic Regression, and Support Vector Machine we see that the sitting, standing, and walking classes were the best performers in classifying because the darker the color of the diagonal entries the better the classification.

Performance Metrics of Classifiers

Classifiers	Accuracy	Precision (PPV)	F1 Score	Sensitivity (Recall)
Naïve Bayes	0.77	0.67	0.64	0.64
Logistic Regression	0.83	0.80	0.77	0.75
Support Vector Machine	0.85	0.83	0.79	0.77
Classifiers	Specificity	NPV	FNR	FPR
Naïve Bayes	0.93	0.93	0.36	0.07
Logistic Regression	0.95	0.95	0.25	0.05
Support Vector Machine	0.96	0.96	0.23	0.04

Fig. 3 Average Performance Metrics

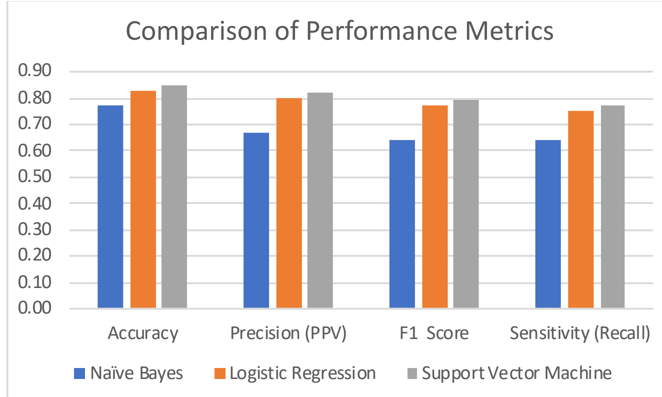

Fig. 4 Average Performance Metrics of Each Classifier

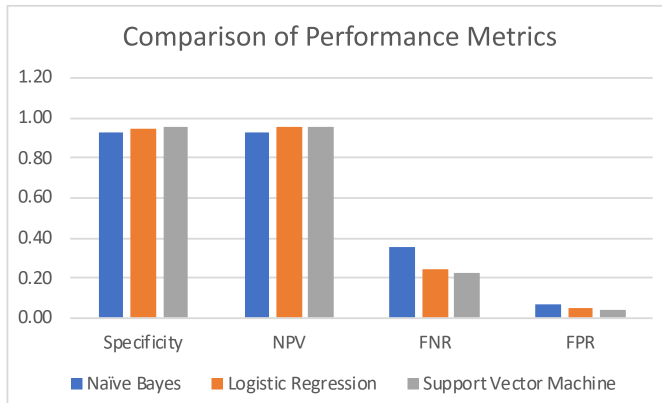

Figure 5 Performance Metrics of Each Classifier

Looking at the calculated performance metrics we see that the accuracy, precision, F1 score, and sensitivity were between 0.64 and 0.85. The specificity and NPV had the highest metrics with values ranging from 0.93 to 0.96 which is very good with 1.0 being the highest metric value. The FNR ranged from 0.23 to 0.36 and the FPR ranged from 0.04 to 0.07 which is very good for FPR since it is close to zero.

IV. Discussion

Comparing my results to [1] the overall performance metrics decreased with the use of less predictors. The average performance metric values that [1] reported was 0.994 for precision and recall, while I had average values ranging from 0.64 to 0.85 a 35% decrease in classification performance. For the FPR my average values were from 0.04 to 0.06 while [1] had an average of 0.001, which shows that with reduction of predictors the FPR increased by 6000%. This suggests that using a reduction of predictors for classifying HAR data the classification performance decreases for the naïve bayes, multinomial logistic regression, and support vector machine with a linear kernel.

V. Future Work & Conclusions

To better improve the classification performance we can explore different classification models such as support vector machines with a radial basis function kernel while varying the parameters cost and gamma, k nearest neighbors, bootstrapping (bagging, boosting, or random forest) by taking replicates of the data with or without replacement, a principal component analysis (PCA) or independent component analysis (ICA) to reduce temporal dimensionality and then apply various classification models.

Acknowledgments

I would like to thank the authors of [1] for providing the HAR dataset to the public.

References

[1] Ugulino, W.; Cardador, D.; Vega, K.; Velloso, E.; Milidiu, R.; Fuks, H. Wearable Computing: Accelerometers' Data Classification of Body Postures and Movements. Proceedings of 21st Brazilian Symposium on Artificial Intelligence. Advances in Artificial Intelligence - SBIA 2012. In: Lecture Notes in Computer Science. , pp. 52-61. Curitiba, PR: Springer Berlin / Heidelberg, 2012. ISBN 978-3-642-34458-9. DOI: 10.1007/978-3-642-34459-6_6. Readmore: http://groupware.les.inf.puc-rio.br/har#dataset#ixzz633AyOfZ4.

[2] Rahrooh, Allen, "Classifying and Predicting Walking Speed From Electroencephalography Data" (2019). *Electronic Theses and Dissertations*. 6281. https://stars.library.ucf.edu/etd/6281