Evaluating Regression and Classifier Metrics

Anuar Konkashbaev

School of Technology & Engineering, National University

Course code: 3602869492

Professor: Dr. Mohamed Nabeel

March 15th, 2025

Evaluating Regression and Classifier Metrics

For this project, I wrote a Python Jupyter Notebook to illustrate the evaluation of regression and classifier metrics for predictive models (Kuhn, 2008). Regression and classification are categories of models or tasks. Regression refers to a model in which the dependent variable is a quantitative continuous variable. Classification implies a dependent variable that is categorical or qualitative. I tested several performance metrics.

For regression, I utilized the mean error (the average of the difference between the actual value and the predicted value), the mean squared error (the average of the squared differences between the actual and the predicted), the mean absolute error (the average of the absolute deviation of the differences between the actual and the predicted), and the mean absolute percent error (the average of the absolute of the percent deviation between the actual and the predicted). For each of these metrics, a lower value implies better performance. For classification, these metrics are not necessarily suitable given the categorical or qualitative nature of the dependent variable. I therefore evaluated the accuracy of the model (the sum of the correctly classified instances divided by the total number of instances), precision (positive predictive value), recall (sensitivity), and the F1-score.

I compared two estimators (est1 and est2) for the “sepal width” variable in cm; see notebook for details. In general, est1 showed better performance than est2. I should point out that I had to modify the est2 estimator because the variable “petal width” from the original assignment does not exist (and hence, I substituted “petal length” to illustrate the metrics). For the classification, I also compared two models (y_hat and y_hat2). The second model, y_hat2, showed improved performance; for example, the F1-score is closer to 1 for y_hat2 relative to the first model, y_hat. In the future, one way of improving the models is to utilize more complex models, such as performing linear regression or multiple regression analysis (Halinski & Feldt, 1970) or incorporating non-linear terms. In multiple regression, with more variables to include as features, one has to consider potential collinearity and consider the interpretation of the model (Mason & Perreault Jr, 1991).

Results:

Results and calculations are publicly available at: https://github.com/anuarAtNcu/DDS8555v1/blob/main/week1.DDS8555v1.assignment1.ipynb

References

Halinski, R. S., & Feldt, L. S. (1970). The selection of variables in multiple regression analysis. *Journal of Educational Measurement*, *7*(3), 151-157.

Kuhn, Max (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software, 28(5), 1–26.* https://doi.org/10.18637/jss.v028.i05

Mason, C. H., & Perreault Jr, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of marketing research*, *28*(3), 268-280.