A22: Accuracy Assessment Metrics with yardstick

In this exercise, you will use the yardstick package to calculate assessment metrics using reference data and model predictions. Deliver the results as an HTML webpage generated from an R Markdown or Quarto file. Use headers or text to differentiate each component of the assignment. Make sure to include both your code and the associated answers.

Multiclass Assessment Metrics: Read in the following files: assessmentR1.csv and assessmentR2.csv. In these files, the pred column represents the numeric code for the predicted class while the ref column represents the numeric code for the reference class.

T1: Recode the pred and ref columns and convert to a factor data type as follows (New → Old): "AnnCrp" = "0", "Frst" = "1", "HrbVg" = "2", "Highwy" = "3", "Indst" = "4", "Pstr" = "5", "PrmCrp" = "6", "Resid" = "7", "Rvr" = "8", "SL" = "9". These data are for a general land cover classification differentiating annual crop, forest, herbaceous vegetation, highway, industrial, pasture, permanent crop, residential, river, and sea/lake classes. (4 Points)
T2: Generate confusion matrices for each result separately. (4 Points)
T3: Calculate the following assessment metrics for each result separately: overall accuracy, macro-averaged class aggregated recall, macro-averaged class aggregated precision, macro-average class aggregated F1-score. (4 Points)
T4: Write a short paragraph comparing the results. Which model generally performed best? Explain sources of confusion between the classes. (4 Points)

Binary Assessment Metrics: Read in the binary_data.csv file. These data represent classification results for differentiating slope failures (coded as "slopeD"), such as landslides, from areas without failures (coded as "not"). The truth field represents the reference class while the predicted field represents the predicted class. prob_not represents the predicted probability for the not slope failure class while prob_fail represents the probability for the slope failure class. These two columns sum to 1.

T5: Calculate the following metrics with the slope failure class as the positive case: recall, precision, negative predictive value, and specificity. (4 Points)
T6: Calculate the area under the receiver operating characteristic (ROC) and precision-recall (PR) curves with the slope failure class treated as the positive case. (4 Points)
T7: Describe the sources of error in the predictions in regards to the relative abundance of false positive and false negative predictions. (4 Points)

Regression Metrics: Read in the regression.csv file. The truth field represents the reference value while the m1, m2, and m3 fields represent predictions made using different methods.

T8: Calculate the root mean square error (RMSE) and R-squared statistics for each model. (4 Points)
T9: Combine the results into a single table. The table should contain a column to differentiate each of the three models, the RMSE results, and the R-squared results. (4 Points)
T10: Based on the results, which model provided the most accurate prediction of the continuous variable. Explain your reasoning. (4 Points)