Math 189: Final Project

Swiss Bank Notes

Working alone, prepare a R Markdown Notebook report based on examining the Swiss bank notes dataset (available from GitHub). The dataset contains six variables measured on 100 genuine and 100 counterfeit old Swiss 1000-franc bank notes:

1. Length of the note
2. Width of the Left-Hand side of the note
3. Width of the Right-Hand side of the note
4. Width of the Bottom Margin
5. Width of the Top Margin
6. Diagonal Length of Printed Area

Analysis

Can we predict whether a note is false or counterfeit using supervised learning?

Your analysis should include the following components for full credit:

1. Explore and visualize the data.
2. Divide into training and validation sets (which each must have some of the genuine and counterfeit notes), implementing K-fold cross-validation.
3. On each fold, classify using both LDA and logistic regression, i.e., fit both models on the training set and evaluate performance on the validation set.
4. Refine the six covariates using a factor model to reduce the dimension and remove any redundancy, re-running analysis on the factor scores.
5. Discuss the assumptions needed (and justify those assumptions, if possible). Does the factor analysis help, or is it a waste of time? Arrive at a final model for each fold.
6. Perform the analysis using R code (if you use a package, it is your responsibility to ensure their code does what you intend it to do).
7. Combine results across folds. Hint: the split used in step 2 will determine modeling results in steps 3 and 4, so cross-validation can be used to determine the best model and any dimension reduction.
8. Summarize the results with appropriate displays (figures and/or tables).

Components to Your Report

This is your report, but consider including the following:

1. Introduction, Body, and conclusion.
2. Data citation.
3. Any relevant tools learned in the course. (However, this project is really focused on lectures 13 through 24.)
4. Justify your analysis.

Ethics

Work alone!!!

Submission

Write an rmd file, render to pdf and submit the pdf.