STAT 574 TAKE-HOME EXAMINATION

(*Due Wednesday, March 15, by 10 PM, sent to stat574s23@gmail.com*)

Do all the exercises below using SAS, SAS Enterprise Miner, R, and Python. Give complete codes (path diagrams) and all relevant outputs.

Problem 1. Use the data in the file "hospital_data.csv" to fit a k-nearest neighbor regression. Compute prediction accuracy within 10%, 15%, and 20% of the actual values. Plot the actual and predicted values in the same coordinate system.

Problem 2. Use the data in the file "card_transdata.csv" to fit a k-nearest neighbor binary classifier with $k = 9$. Compute prediction accuracy.

Problem 3. Use the data in the file "concussions_data.csv" to fit a k-nearest neighbor multinomial classifier. Compute prediction accuracy.

Problem 4. Use the data in the file "hospital_data.csv" to fit a support vector regression with linear, polynomial, radial, and sigmoid kernels. Compute prediction accuracy within 10%, 15%, and 20% of the actual values. Choose the best-fitted model. Use R and Python only.

Problem 5. Use the data in the file "card_transdata.csv" to fit a support vector binary classifier. Specify linear, polynomial, radial, and sigmoid kernels (whichever are possible to fit). Compute and compare prediction accuracies.

Problem 6. Use the data in the file "concussions_data.csv" to fit a support vector multinomial classifier. Specify linear, polynomial, radial, and sigmoid kernels (whichever are possible to fit). Compute and compare prediction accuracies. Compute prediction accuracy.