Rensselaer

why not change the world?®

Lab 4 : Principal Component Analysis

Ahmed Eleish

ITWS-4600/6600/CSCI-4600,

October 24th, 2025

Tetherless World Constellation
Rensselaer Polytechnic Institute

Lab 4

https://rpi.box.com/s/76tfin2br8gl1nooubt47ymg72zw47l3

Principal Component Analysis

Using the wine dataset:

- Compute the PCs and plot the dataset using the 1st and 2nd PCs.
- Identify the variables that contribute the most to the 1st PC and 2nd PCs.
- Train a classifier model (e.g. kNN) to predict the wine type using all the variables in the original dataset.
- Train a classifier model to predict the wine type using the data projected onto the first 2 PCs (scores in the princomp function's return object)
- Compare the 2 classification models using contingency tables and precision/recall/F1 metrics.

Please push to your github repository:

1. All your code in a *.R or *.MD file
2. All outputs (plot + contingency tables)

Thanks!
Have a great weekend!