

Problem n.1

Eddie and Dustins are interested in analysing the characteristics of the Dungeons & Dragons monsters. The files `dnd_monsters.txt` contains 401 monsters with their characteristics:

`armor.class`: how difficult the creature is to hit
`hit.points`: how much damage the creature must take before being defeated
`strength`: how easily can the creature smash a tomato
`dexterity`: how well could a creature avoid getting hit with a tomato
`constitution`: could the creature eat a moldy tomato and not get sick
`intelligence`: does the creature know if a tomato is a vegetable or a fruit
`wisdom`: does the creature know what to pair a tomato with
`charisma`: could the creature sell you a tomato
`size`: how much space it occupies (categorical).

- a) Perform a Principal Component Analysis of the dataset, by only focusing on the quantitative variables of the dataset; here, evaluate whether it is appropriate to use the original variables or the standardized ones and proceed accordingly.
- b) Report a plot of the loadings of the first two principal components and provide an interpretation.
- c) Report the scatter plot of the data along the first two PCs. Use the categorical variable `size` to interpret the results.
- d) Consider now only the monsters of dimension `Tiny` and `Huge`. Use a support vector machine with linear kernel and cost equal to 1 to classify the monsters. Report the number of support vectors. Plot of the classification regions. How would you classify a monster with `armor.class=14`, `hit.points=50`, `strength=19`, `dexterity=10`, `constitution=16`, `intelligence=8`, `wisdom=12` and `charisma=13`?

Upload your results here:

<https://forms.office.com/Pages/ResponsePage.aspx?id=K3EXCvNtXUKAjjCd8ope6-9ASOGwf2lHjvGX24HiqFVUQUNHSUo2R0YxVTFRT1BDM1lJTEtDWExGUy4u>