Knowledge-Based Systems in Bioinformatics - project tasks

1. **Create a rule-based classification model**
   a) Use rmcfs to identify significant features that would be best to use for classification.
   b) Use R.ROSETTA to create and train a rule-based model using the selected features. You are encouraged to try different algorithms for discretization, reduct calculations and other computations. Don't forget to estimate the performance of the classifier.
   c) Create VisuNet graphs for the constructed models

2. **Prepare a presentation**
   Make a short presentation (approx. 20 minutes) about what you have done. Use the following questions to plan your presentation:

a) What was the background of the data?

b) What kind of the data was used? How was the data obtained? Specify quality and quantity of the dataset.

c) What were the most important features?

d) What was the aim of the classification model?

e) What kind of models and algorithms did you use? Explain why.

f) Which parameters did you choose? Explain why.

g) How did you evaluate your models?

h) Give an interpretation of the VisuNet graphs. Show the strongest connections, hubs or separations etc.

i) Report the top rules for each class and provide a brief interpretation of them.

j) Report the features that discern between different classes (from the top 5 rules).

k) Investigate the top rules for common features. Compare the results with the VisuNet network.

l) What results did you get? What was the performance? Which methods worked best?

m) Were the models themselves informative? Give an example of what you can learn from the models.

n) What was difficult or could be improved in the project?

Hints

- **rmcfs**
  Please note that the agreement between the R and Java version is required i.e. 64-bit or 32-bit.

  UNIX
  In case of any problem with the rJava package, run the following command and then reopen RStudio: `sudo R CMD javareconf`

- **R.ROSETTA**

  The package is publicly available on GitHub:
  https://github.com/mategarb/R.ROSETTA
  The tutorials are available on the website:
  https://komorowskilab.github.io/R.ROSETTA

- **VisuNet**

  The package is publicly available on GitHub:
  https://github.com/komorowskilab/VisuNet
  The documentation is available on the website:
  https://komorowskilab.github.io/VisuNet/