

Questions:

What is the research question?

The aim of this study was to compare the performance of three feature selection methods, named ReliefF, SVM-RFE and Information. So there research question was : What feature selection methods is best suited array-CGH breast cancer data

What is the validation scheme?

She validates using the k-fold cross validation. The ratio between training and validation is 75:25

What is the main result? / What is the answer to the research question?

There is only one result : RFE with 50 features seemed perform best. This is also the answer to there research question which feature selection method would be best.

Feedback:

Global feedback :

- 1) This paper should be written in the form of a Manuscript. You need to use the words or latex template provided see link:
http://www.oxfordjournals.org/our_journals/bioinformatics/for_authors/submission_online.html
- 2) The reference should start with one and continue on, and not with a six. This makes it harder for a reader to be able to use your references to look up a reference.

Research question: 4

The research question is clear and can be answered with the data given. You are focusing on 3 feature selection methods but like you mentioned every ML method will have a different preference for feature selection methods so you maybe need to choose a method and then try and find the feature selection methods that best suite them because not only the type of data will make a difference but also the right ML method for prediction.

Context: 3

There is a biological connection to the research question but it is a bit harder to find in your last sentence you mentioned: optimization of array-CGH data analysis, therefore improving breast cancer classification and its treatment. But you don't mention how this would improve or why this is so important to perform classifications and, thus being able to separate the 3 sub-types and improved prognosis and treatment.

Methodology: 2 .5

The feature selection methods are explained but are not clear. You should be able to understand how those methods work from the text . But you only mentioned what the method does so you need to elaborate a bid in these sections especially since this is really important for your research question. Training/CV scheme mentioned in both a figure and explained in the text. It does make sense but if you improve the workflow this will be a lot better.

Contents/Results: 1

The results section consist of 1 table so a lot is missing.

Tables/Figures: 2

The The workflow is not very clear. The figure contains a lot of abbreviations

that were not obvious and the overall flow was hard to follow. Figure 2 is not really useful we are only allowed to put 4 figures in the paper and this only explained a classification method so if you explain the method more this would not be necessary.