KTH-CSC/DD142X

Peer review protocol

Authors: Martin Hyberg and Teodor Hurtigh Isaacs
Report: Sentiment Analysis
Reviewer: Robin Gunning <Rgunning@kth.se>
Date: May 23, 2018

1 Summary of the report

The authors have studied sentiment analysis, which in short is to classify a piece of text as either positive or negative. This is supposed to be done on 50 different youtube videos chosen based on specific criteria which is not presented at the time of this draft. The comments on the youtube-video will be used to ”predict” if the ”like-ratio” will be positive or negative.

The authors presents a few different ways to handle this task but there is no section yet explaining which method the authors will use to get their results.

The authors have a placeholder image in the result section but there is not much information about what the results mean. My guess is that the result is non conclusive and you can't really predict the ”like-ratio”

2 Title and abstract

The title is not very informative and seems like it is incomplete. Sentiment Analysis could mean anything. At this point there is no abstract to comment on.

3 Introduction and problem statement

The introduction is fine, but some wording problems exist such as:

> ”Thanks to social media everyone with a network connection has the opportunity to voice their opinions at a global scale regarding current events worldwide".

Global and world wide have essentially the same meaning.
The problem statement is too short and just scarecly touches on the problem at hand, the problem statement should contain the use of machine learning or it might aswell ask if someone manually can guess the like-ratio from reading the comments

4 Background and related work

4.1 Brain Computer Interface (BCI)

The authors write:

> In this thesis this will be done using three different machine learning algorithms to classify each comment into two categories, positive or negative.

If this is the case for BCI in general then your task seems somewhat contradictory. Generalizing a system to work cross-individual, but the system requires the new individual to learn the system? Maybe you intended to say that this is historically the case for (non-cross-subject) BCI?

4.2 Neural Networks (ANN)

Perhaps the section should be named "Artificial Neural Networks" to match the given ANN acronym. Moreover, the section feels extremely abstract and non-technical. Perhaps you could include a typical textbook example of a basic ANN.

Section numbering would be helpful in showing that the following sections on CNN and RNN/LSTM are actually subsection to the ANN section.

5 Methods

5.1 Data

I'm not sure I know what is meant by "The data was preprocessed using absolute fast fourier transform". Perhaps include some background on EEG data and its (pre)processing. Maybe this is something that is painfully obvious to someone familiar with processing EEG data (which I am not), in which case I guess it could be left as is.

5.2 Model selection - SVM

The following sentence is strange: "In the input data for the models there are 16 frames with summarized samples for beta and mu for the C3 and C4 channels respectively there are in total 64 features at this point."

5.3 Model selection - LSTM

You write ”For creating the LSTM we used Keras [...]”. You could perhaps mention this ”Keras” in the background or atleast include a footnote with a link. Including specific version numbers might also be appropriate.

5.4 Model selection - CNN

You write ”Similarly the CNN model was created [...]”. Similarly to what?

6 Results

Maybe I missed it, but I don’t remember seeing a description of how you selected data for / trained the ”model trained on all data” in the Method chapter.

The descriptions of ANOVA and hypothesis testing found at the end of the Results chapter should be moved to the Method chapter. ANOVA should be capitalized.

6.1 SVM

The number of digits in ”Accuracy of model trained on all subjects 0.5251141552511416” seems excessive.

7 Discussion and conclusions

7.1 Results

You write:

> [...] three [sic] seemed to be a trend of high inter-subject and inter-session accuracy for subject s1, s2, low for subject s3 and s4. This could be taken to imply that some individuals produce EEG data more suitable for generalization than others and vice versa [...]

Besides the initial typo (”three”, should be ”there”), this seems to me a strange way of thinking about things. If the models can generalize over certain groups of people but not others, then surely it is the case that the models are not good enough for achieving a general solution, not that some individuals just produce unsuitable data. It seems more reasonable to say something like ”This could be taken to imply that there exists subsets of people that certain models are able to generalize over”.

8 Overall characteristics

Disposition structure and language are both fine. Coherence is only slightly broken in a couple of places where ambiguous or undefined terms are used; e.g the first uses of the terms ”performance” and ”accuracy”, the use of ”single-subject BCI classifier”, the use of ”Keras”, potentially the use of ”extracranial” depending on the intended audience. Overall coherence is fine.

The style could be improved by adding chapter and section numbering and by printing page numbers on each page. Tables in the Method and Results chapters should be numbered. Graphs in the results chapter should be made into numbered figures. Figures 3, 4 and 5 seem not to be referred to in the text.

9 Overall impression

The strongest points are structure and language. The paper is easy to read and to follow, meaning the essential parts are all mostly present. The weakest points are the unnumbered figures and tables.

10 Suggestions

These are the suggestions made throughout this document, collected in a bullet point list for convenience.

- Style
  - Add chapter and section numbering.
  - Print page number on each page.
  - After numbering all figures (e.g graphs) and titling + numbering all tables, make sure to reference each figure and each table in the text atleast once.
  - Capitalize ANOVA.
- Abstract
  - Re-formulate the statement ”Our results show that inter-subject generalization depends on the subject”.
  - Name the factors affecting generalization ability and your suggested future avenues of research.

- Introduction and problem statement
  - Clearly define the term ”single-subject classifier”.
  - Clearly define the terms ”performance” and ”accuracy” in your context.
  - Consider adapting the problem statement to the results, it seems you are not really investigating the stated specific relationship.
- Background
  - Re-formulate or expand the paragraph on how users are required to learn to control a given BCI such that the idea of a generalized system becomes less contradictory.
  - Rename the section ”Neural Networks (ANN)” to ”Artificial Neural Networks (ANN)” such that the acronym matches.
  - Add some brief technical details on ANNs.
  - Add a section on EEG data and its preprocessing using absolute fast fourier transform.
- Method
  - Number and title all tables.
  - Include a clear description of how the data was selected for the ”model trained on all data”.
  - Fix a strange sentence (see 5.2)
  - Include a footnote with a link to Keras
  - Re-formulate the leading sentence in the section on CNN model selection (see 5.4)
- Results
  - Have all graphs be numbered figures.
  - Number and title all tables.
  - Move the descriptions of ANOVA and hypothesis testing to the Method chapter.
  - Re-formulate the leading sentence in the section on CNN model selection (see 5.4)
  - Reduce or motivate the large number of digits in the result for the SVM model trained on all subjects.

- Discussion and conclusion
  - Fix a typo ”three seemed” -> ”there seemed”.
  - Consider re-formulating the parts on how some people seem to generate unsuitable data (see 7.1)
  - Reduce or motivate the large number of digits in the result for the SVM model trained on all subjects.