

- Summary of the report.

Sentiment analysis is the topic which this report focuses on. Reviews in Chinese language from Weibo social network are the data source, and the project aims to classify those as positive or negative. Several approaches from data aspect and model aspect have been proposed, as reviews are translated to English, or contracted to Emoji only to examine the performance of 3 models, varying from classical methods KNN, Naïve Bayes to state-of-the-art NLU model BERT. The result shows that figuring out the core feature of data is very important, say in this case, emoji almost always represents the true sentiment of reviewers

- Describe the strengths of the report.

It proposes an interesting approach (translation), explore a wide range of ML technique. Also, it shows an adequate concentration on data.

- Describe the weaknesses of the report.

Just a minor thing to say: TF-IDF statistic is introduced but not applied, making confusion. The expression corresponding to value $p = 0.2$ is not “norm”, as p needs to be ≥ 1 to call the expression norm. Bert needs to be CAPITALIZED as BERT.

- Evaluation on quality of writing (1-5): Is the report clearly written? Is there a good use of examples and figures? Is it well organized? Are there problems with style and grammar? Are there issues with typos, formatting, references, etc.? Please make suggestions to improve the clarity of the paper, and provide details of typos. (4)

The overall organization is fine, easy to follow.

- Evaluation on presentation (1-5): Is the presentation clear and well organized? Are the language flow fluent and persuasive? Are the slides clear and well elaborated? Please make suggestions to improve the presentation. (4)

They go a bit fast on their own KNN so that there is not enough press on the difference between their KNN and usual KNN. In class, BERT has been introduced, so they don't need to talk about that so much.

- Evaluation on creativity (1-5): Does the work propose any genuinely new ideas? Is this a work that you are eager to read and cite? Does it contain some state-of-the-art results? As a reviewer you should try to assess whether the ideas are truly new and creative. Novel combinations, adaptations or extensions of existing ideas are also valuable. (4)

- Confidence on your assessment (1-3) (3- I have carefully read the paper and checked the results, 2- I just browse the paper without checking the details, 1- My assessment can be wrong) (2)