Mike Dean

Peer Review of Fianna O’Brien

1. Summary: Start writing a short summary of the paper (3-5 sentences) that describe the main points raised in this paper/project.

This paper was a summary of a statistical analysis of internet message board comments in order to classify them based on their negativity or neutrality. Fianna used Python’s Natural Language Processing library to analyze the data and perform machine learning on it. Additionally, the difficulties surrounding natural language processing in machine learning are discussed.

1. Rubric of quality of work: Fill out the rubric on the following page indicating the quality of each topic listed (scale 1-4). You can list additional comments related to the rubric entries as well.

1. Point-by-point comments: Make a list of particular points in the paper/project that could be improved upon. Are there parts that are unclear or where the organization of the paper could be better? Are there additional topics that should be looked into on a data analysis project?

I only came across a few grammatical errors, most likely typos. I think the paper would really benefit from a little bit of insight into the ramifications of cyber-bullying. It mentions that most young adults have experienced it, but not how it has affected them. Finally, I think that the code included at the end should be commented. The code is formatted well, so you could just add a comment at the beginning of each block briefly describing what each function does so the reader can more easily follow it.

1. Overall impression: In a typical journal article one would state whether a paper should be accepted, rejected with minor revisions (that the author can fix to get it accepted), rejected with major revisions (that the author can fix to get it accepted), or rejected without the opportunity to revise. Here we obviously won’t reject each other’s work, but you should state your overall impression of the quality of the work, and what can be done to improve the paper/project.

I found this paper to be very well written and interesting. I think that if you fix the few grammatical errors and add comments to your included code, it would be perfect.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Excellent (4) | Good (3) | Fair (2) | Inadequate (1) | Rating |
| **Relevant: Is the paper/project about an interesting prediction problem?** | The paper/project covers an exciting prediction problem that is clearly related to the class | A good/interesting prediction problem related to the class | It is questionable whether the project discussed is that strongly related to the class | The topic is only tangentially related to prediction | 4 |
| **Organization: Do the**  **topics/analyses follow a logical progression?** | The paper is well organized describing why the problem is interesting, citing relevant sources to back up key points, and follows a logical progression of topics. | The organization of the paper is good, although more sources could be cited to back up key points. | The paper has a logical flow to it, although it is not clear why certain topics were covered and more sources are needed. | The paper is not well organized, it is unclear why certain topics are raised, and ideas are stated without sources to back them up. | 3 |
| **Quality of writing/analyses: Is the paper well written (analyses well done) with few errors?** | Clearly describes the problem, the methods that were used to make the predictions, and a critique of the methods. Overall enjoyable to read. | Describes the problem, the methods that were used to make the predictions, and a critique of the methods, although the writing could be clearer in certain places. | The writing could use some improvement in certain places. | The writing is not clear. | 4 |
| **Citations\*: Are there a**  **significant number of reputable sources cited to backup claims, and are the citations in a consistent style?** | High quality (scientific) sources are used which strongly reinforce the points made in the paper. The citations followed a standard convention. | Scientific/reputable sources are used to reinforce the points made in the paper. The citations followed a standard convention. | Not all sources are reputable. The citation style  is not fully consistent. | Many sources are reputable. The citations do not follow a conventional style. | 4 |
| **Analyses\*\*: The analyses done give insight into answering a prediction question** | Yes, the analyses are interesting and give clear insight into a prediction problem. | The author is just learning the analysis methods, so the predictions are not yet accurate, but the methods seem sound. | The author tried out a few methods but does not yet seem to fully understand the technique. | The analysis are weak and it is not clear that much effort was put in to understand the methods used. | 4 |

\* Most relevant for papers. \*\* Only relevant for data analysis projects.