|  |  |  |
| --- | --- | --- |
| **Section** | **Comments** | **Justification** |
| Introduction | This section is pretty good. Please make sure you used all the most cited references on previous studies using social media data to investigate infrastructure and community during disasters.  The only paragraph which is not really clear in the second last. Please try to make it easy to read making the text simpler. **(Full section)** |  |
| This is a ref from 2017 it is a bit old to say how many tweet are used today. Better integrating with a newer reference **(4th para)** |  |
| Please add a good reference supporting this part **(4th para)** |  |
| I am not sure what you are trying to say here. Please simplify **(5th para)** |  |
| I do not understand how this is linked with the previous part. Please simplify. **(5th para)** |  |
| Literature review | This section needs to be rewritten as it does not read a literature review. Given the goal of the work you need to provide a review on how social media date have been used to investigate disasters. You can cite the most important works and then existing literature reviews.  Then you zoom in on the use of Twitter data and how they have been used so far. Try to find all the existing studies doing something similar to what you do in this paper. Explain advantages and limitations of existing works.  Finally explain why we need to propose a new solutions given the research gap identified in the review and say that the new solution is proposed in the following section **(Full section)** |  |
| Please try to use simpler words. I don't know what this word means **(1st para)** |  |
| Background: The Kerala Floods | Better provide a map showing the area under investigation through a figure. See for instance the paper I wrote on Brindisi as an example.  Make sure to highlight in the map the areas affected by the flood. See the examples of my wildfire papers.  You need a timeline describing the flood you are investigating here. When it started according to official records and how it evolved and major events during the flood **(1st para)** |  |
| Please add the equivalent in US dollars in parenthesis **(1st para)** |  |
| Framework | Please add the references of all the python and R reference you used in this work. Any external piece of code need to be referenced here |  |
| Figure should go after the text mentioning them.  This figure can improve providing the information of the libraries used and if they are in python or R **(Fig 1)** |  |
| 3.1. Data extraction | Probably better move all the code in a final appendix instead of keeping it in the paper |  |
| 3.2 Removing duplication | Is there a figure or equation that can explain how it works?  We need to give a sense of what it does for people who are not expert about it like me |  |
| 3.4 Data wrangling | Beautiful idea. You could provide an example for all of the 6 pre processing modules |  |
| Same as for the previous code **(Fig 4)** |  |
| Results | This section still provide information on the framework instead of the results on how the framework is applied.  The results show provide information on how many raw tweets you used and how many you have after the pre-processing then you start showing the results on geo-tagging, distance distribution etc.  Any technical information on API and coding used here should go in the previous paragraph Framework **(Full section)** |  |
| Now it is pretty simple visualization. Can you show how these charts change over time during the different stage of the flood?  This can make the paper stronger **(Fig 5)** |  |
| I cannot read numbers and text in it. It would be good to put information on when the flood started and major vent in the flood (the one you should mention in the Background section) **(Fig 6)** |  |
| Now it is pretty simple visualization. Can you show how these charts change over time during the different stage of the flood?  This can make the paper stronger **(Fig 7)** |  |
| Please improve the quality of these charts. I cannot read what is written in it.  Now it is pretty simple visualization. Can you show how these charts change over time during the different stage of the flood?  This can make the paper stronger **(Fig 9)** |  |
| Discussion | You need a good discussion section here  First you need to explain what is new in your proposed solution and how it is different from what you have show in the literature review section.  Then you explain what you have learn applying the proposed solution to the specific case study  Finally you start talking about limitations and what can be done in the future to improve it  **(Full section)** |  |