Review 2

1. Paper id: *

27

2. Author(s): *

Crisan Dragos, Coroiu Adriana

3. Title: *

Implementation of machine learning algorithms in precision agriculture using soil data

4. General evaluation: *

- ( ) Very good - contains original work
- (•) Good - not original work, but methodologically sound, detailed
- ( ) Acceptable - correct work, but could have been improved
- ( ) Poor - unacceptable

5. Overall recommendation: *

- ( ) Accept as it is
- ( ) Accept with minor modifications
- (•) Weak accept (major modifications are needed)
- ( ) Reject

6.

For each criterion, please use the scale to answer the question:
"***To what extent does the article meet this criterion?***" *

	1. Fails by a large amount	2. Fails by a small amount	3. Neutral	4. Succeeds by a small amount	5. Succeeds by a large amount
The subject addressed in this article is worthy of investigation.	○	○	○	○	◉
The paper contains original work.	○	◉	○	○	○
The title of the paper reflects sufficiently and clearly the topic.	○	○	○	○	◉
The abstract contains a sufficient summary of the work.	○	○	○	◉	○
The introduction is relevant to the topic and has clear objectives.	◉	○	○	○	○
The research method is solid and proven.	○	○	○	◉	○
The results are clearly presented with sufficient analysis.	○	○	○	◉	○
The conclusions are supported by the data	○	○	○	◉	○
The references are adequate and appropriate.	○	◉	○	○	○
The English is satisfactory.	○	○	◉	○	○

7. Is this paper a good candidate (appropriate) for publishing in Studia Informatica?

(Please do not mark with YES papers that are highly likely to be rejected immediately by the reviewing panel of Studia Informatica – because of their quality. You are strongly suggested to propose the required modifications to make it a good candidate for publication)
*

○ Yes

◉ No

8. If YES – please score the quality of this paper from the perspective of publication in Studia Informatica:

1	2	3	4	5

9.

Remarks concerning the contents and recommendations for optional or compulsory modifications. Please give a frank account of the strengths and weaknesses of the article: *

The authors present some experiments performed for predicting the crop yield. They have build models using XGBoost, CNN and FCNN and compared their performances. The experiments showed that XGBoost performs best with a really high accuracy.

Below you have a long list of observations. In conclusion the research is interesting and seems correct, but there are many modifications which need to be done to make this paper look like a scientific paper. It should be restructured and in some cases more information should be added.

- In the Abstract you mention "This model was trained on a public dataset situated in India" - I don't think the data set is in India, but rather the data was collected in India.
- The Abstract should contain 1-2 sentences about the performance of the proposed approaches as well.
- The Introduction needs to be rewritten completely, since currently it sounds too much like a few paragraphs taken directly from your bachelor/dissertation thesis:
- The first sentence of the Introduction is problematic: first, there are 3 Sections, not chapters and such a description of the structure of the paper is in general the last paragraph of the Introduction. It sounds really strange to begin with something like that. And you do not need to write "paper" with capital P.
- The goal of introduction is to present the approached problem, to explain why it is important, to mention what original contributions the paper brings. None of these is currently present in the Introduction.
- Short forms (ex. can't, don't) are not used in formal text.
- In the last paragraph of Section II, it should be "its" not "it's" (this problem appears in multiple places in the paper).
- At the end of paragraph 2 from section II: "by the raising concerns of food security" it should be "rising concerns". In the same paragraph you have R^2 = 96.33. It is either 96.33% (as percentage) or it should be between 0 and 1 (as it is in the next paragraph).
- In section II, where you talk about the methods used in paper [2], the "vector machine" isn't actually "support vector machine"? Because I do not know what "vector machine" is.
- The last paragraph if Section II should not be there, it is not about Related Work. Maybe, it should be better suited in Introduction (with some modification) or in a section discussing the results.
- First sentence of III.A.1) should be rewritten: as I mentioned, I do not think that the dataset itself is in India and it (the dataset) certainly was not conducted in the years 2018-2019. I do not know the dataset, but I guess the author meant to say something like "the dataset contains data collected from India in the years 2018-2019".
- In the description of the dataset all 4 parts should be mentioned and only after the author should describe which part of the dataset they used. Also, it is strange that the first sentence says "it is split in 4" and then "three datasets were merged". I checked the dataset: it contains 4 files (and you should shortly present all 4 of them) out of which the author used 3 (explain which 3) merged together. As mentioned, I think the attributes, the structure, number of instances for the datasets should be presented before presenting the data transformation and analysis.
- When you reference a Figure or table you need to use the word "Table", "Figure", not just the number, which in itself looks really strange, especially for Figure 1, where it is just (1).
- Text on the charts from Figure 1 is not really readable. I know that it is not easy to put so many charts on a Figure, but then maybe more explanations should be given about the content of the Figure. Also, what i the benefit of having both a histogram and a box-plot of the same information?
- In the descriptions of Figure 2, instead of "gradient", I would say "intensity" or "shade".