

# Review of: "Application of Data Mining Combined with K-means Clustering Algorithm in Enterprises' Risk Audit"

Leila Zemmouchi-Ghomari<sup>1</sup>

<sup>1</sup> École Nationale Supérieure de Technologie

**Potential competing interests:** No potential competing interests to declare.

This work is about using DM techniques, in particular combining the K-means algorithm and RF-LightGBM to classify the risk audit of e-commerce enterprises. It has been applied to the Jingdong Mall dataset.

The obtained results are promising and performed better than other algorithms.

However, the paper is not well written.

Several parts of the paper need to be improved. I cannot detail all the remarks here; this is why I uploaded the paper with annotations on the parts that need to be justified, clarified, or reorganized. To summarize:

The application domain, the risk audit of e-commerce enterprises, is not well described. We have no idea about the relevant features for classification prediction, nor about the obtained classes outputs.

Related works are just described; highlights on their limitations or challenges are not presented, which motivates the current work.

The definitions of the used algorithms should be moved to a background section or to an appendix. A clear separation should be made between the theoretical background and the case study. Because of this mix, the paper is difficult to read.

The case study must be detailed to enable readers or researchers to reproduce this work.

Most of the figures are useless; they are generic and not specific to the case study. The main algorithm that describes the process of combining the two algorithms is not explained. The dataset used for experimentation is not described.

In conclusion, although the work is interesting, major revisions are required before publication.