Task 3

MIDTERM ASSESSMENT

Student ID:20301393

Student Name: Fahim Ahamed Romit

Course Code: CSE449

Section:01

Review: Assessing Trustworthy AI in Times of COVID-19: Deep Learning for Predicting a Multiregional Score Conveying the Degree of Lung Compromise in COVID-19 Patients
URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9845195

1 Summary of the paper

1.1 Motivation/purpose/aims/hypothesis (1 pt)

The paper is against this background the paper focuses on the need for trustworthy AI in healthcare during the COVID-19 pandemic. Authors seek to present work done towards assessing the trustworthiness of an AI system which predicts a multiregional score, known as a Brixia score, from CXR describing the degree of lung compromise in COVID-19 patients.

1.2 Contribution (1 pt)

The AI system proposed will provide the automated framework to assess the severity of lung damage in COVID-19 patients using the CXR modality. It also assists the radiologist in faster and more consistent diagnostics that are explainable to arrive at optimal decisions during the ongoing pandemic.

1.3 Methodology (2 pts)

The Brixia score was to be forecasted using the BS-Net deep learning system in this methodology. It had segmentation alignment, feature extraction, and explainability mechanisms incorporated into the model. The assessment of ethical and technical, medical consideration for trustworthiness within the AI system was done using especially sociotechnical scenarios by using the Z-Inspection methodology

1.4 Conclusion (1 pt)

The paper concludes that, with the promise of support that overworked radiologists provided during the pandemic, the AI system requires further testing and refinement to be generalized outside of the particular dataset and hospital conditions.

2 Critiques or limitations

2.1 1st Critique/Limitation (1 pt)

The AI system was trained with a data set that is not diverse since most of the patients came from just one hospital in Italy; this might result in the lack of generality of the system when applied to diverse populations.

2.2 2nd Critique/Limitation (1 pt)

The Brixia score, which the system is based on, has a semi-quantitative nature; further, inter-individual variability among radiologists introduces variability in developing a standardized and universally applicable AI assessment tool

2.3 3rd Critique/Limitation (1 pt)

It is where this becomes limiting, as the AI model is specific to COVID-19 lung damage, while the Brixia score is not limited only to COVID-19, nor does it consider other pathologies/infections such as co-infections.

3. Synthesis (2 pts): Possible Applications/Future directions

3.1 1st potential/idea of a new/follow-up/extension paper

This could be extended by training the model with more diversity in its dataset, including but not limited to patients from various geographical regions and ethnic backgrounds, which may show how well the AI generalizes across different populations.

3.2 2nd potential/idea of a new/follow-up/extension paper

The second alternative would be to add the capability to the AI system to assess other lung disorders aside from COVID-19. That would widen the application of this kind of tool for more clinical purposes related to the detection and monitoring of lung diseases such as COPD, pneumonia, and tuberculosis.