Article Reviewed:

A. Holzinger, G.Langs, H. Denk, K. Zatloukal, and H. Muller, "Causability and explainability of artificial intelligence in medicine", Wiley Interdiscip Rev Data Min Knowl Discov. ,2019, Vol. 9, No. 4, 10.1002/widm.1312

The article presents the concept of causability and explainability of artificial intelligence in medical field. It mainly focuses on the need of causability for interpretation. To have an effective trustworthy AI system there is a need of causability other than explainability of AI. Explainability is, using what algorithms and steps the AI system reached to a specific decision while Causability demonstrates the quality of the explanation understandable by the human i.e. what are the reasons that a certain decision is made by the system. It thoroughly explains the concept of explainability of AI in decision making. Among the three types of explanations, peer-to-peer explanation is emphasized. Causability would help the expert to understand and learn why a certain decision is made, to correct the errors and to imitate when required. There are two types of explainable approaches, posthoc and antehoc. Posthoc explanation is after the occurrence of an event and antehoc explanation is before an event occurs. Explainability is further elaborated by the example of interpreting a deep neural network and the human explanation of use-case of histopathology. Using AI to support medical decision, it is necessary to understand how and why a decision is made.

In the abstract, author has stated the argument that with explainability of AI system causability is also needed. However, the abstract fails to explain what the paper is about. You must go through the whole article to know that author is aiming to propose causability a new solid scientific field. Moreover, why the use-case of deep learning interpretation and human explanation in histopathology are provided is not clearly mentioned. These examples depict the complexity as well as demand of causability in a system.

The language of the article is effective and easily understandable. It has logical and justifiable explanations of the technical terms in field of Explainable AI. The author has precisely summarized the concepts that are readily comprehensible by any reader.

The article gives a new direction to the researchers and developers in the field of AI however this is not a paramount contribution to Explainable AI. The author has been mainly focused on the issue but the solution to the problem is overlooked. It is fairly concluded that AI systems lack explicit declarative representation of knowledge therefore for an efficient, effective and user satisfied AI system, the factor of causability should be included but how causability could be incorporated is not discussed.

Furthermore, it can be inferred from the abstract that the paper is about need of causability in explainable AI, but as you read further the paper is mainly emphasizing explainability. It is evident that most of the research and facts are present in the domain of Explainable AI but there should be more counter statements as to why causability is required. The examples make evident the nature of complexity of explainability in both machine and human decision. After mentioning the interpreted definition of the causability, the author has not discussed about it until in section 4.1, to purpose the research direction in structural causal models but there is still no methodology. The article contains theoretical references rather than having logical reasoning or technical methods to justify the need of causality.

Overall, the article highlights the issue that already exists and does not provide a novel solution to the problem. It is based on the literature review and does not includes any solid research by the author. But it could help in understanding what the field of artificial intelligence is aiming for.