Article reviewed:

Fan, Xiuyi, Siyuan Liu, and Thomas C. Henderson. "Explainable AI for Classification using Probabilistic Logic Inference." ***arXiv preprint arXiv:2005.02074*** **(2020).**

The article presents an explainable classification approach that produces accurate explanation and prediction of behaviour of an AI system. The classification is solved as inference with linear programming on a probabilistic Knowledge Base learned from the training data. It is demonstrated that whether a query instance is positive and if so, what features characterizes it to that result. The Knowledges bases are set of clauses and probability of those clauses. For the construction of Knowledge Base two methods are presented, one to construct from a decision tree and second to construct directly from the training data. Binary classification problem is considered to illustrate the methods. KB construction method produces clauses with probabilities where a clause is a disjunction of literals and literal is a propositional variable or its negation. A query instance asserts some feature-value pairs. For the inference method, the probability of positivity under these feature-value pairs is calculated. If it is greater than 0.5 than the query instance is categorized as positive. An algorithm is proposed to improve query efficiency. Using that a knowledge base is constructed containing only query relevant clauses. The algorithm is tested on synthetic and non-synthetic data sets giving satisfactory results comparable to existing algorithms thus validating the approach.

The language of the article is incomprehensible. Basic probabilistic concepts used to define the statements are explained complexly. It takes three to four reads to understand the meaning expected to be delivered. Moreover, the information is cluttered. It is difficult to focus due to low maintenance of information density. The algorithm 5 presented in the article is not analysed with the practical data sets thus contradicting the improvement of query efficiency by this method.

However, the article is contributing to explainable AI by proposing a non-parametric classification technique using probabilistic logic inference. The advantage is that the classification method supports partial queries. Since the inference process tolerates inconsistency of knowledge base, incomplete or imperfect knowledge can be incorporated. On comparing with state of the art Shapley Value based explanation method, the approach shows superior results when tested on four synthetic data sets. Comparing to the NonlInear Probabilistic Logic Solver (NILS), this approach tolerates inconsistency.

Overall, it a good research article, introducing a better approach for classification in the field of Explainable AI.