World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:19, No:08, 2025

The Proposal for a Framework to Face Opacity and Discrimination 'Sins' Caused by Consumer Creditworthiness Machines in the EU

Authors : Diogo José Morgado Rebelo, Francisco António Carneiro Pacheco de Andrade, Paulo Jorge Freitas de Oliveira Novais

Abstract : Not everything in AI-power consumer credit scoring turns out to be a wonder. When using AI in Creditworthiness Assessment (CWA), opacity and unfairness 'sins' must be considered to the task be deemed Responsible. AI software is not always 100% accurate, which can lead to misclassification. Discrimination of some groups can be exponentiated. A hetero personalized identity can be imposed on the individual(s) affected. Also, autonomous CWA sometimes lacks transparency when using black box models. However, for this intended purpose, human analysts 'on-the-loop' might not be the best remedy consumers are looking for in credit. This study seeks to explore the legality of implementing a Multi-Agent System (MAS) framework in consumer CWA to ensure compliance with the regulation outlined in Article 14(4) of the Proposal for an Artificial Intelligence Act (AIA), dated 21 April 2021 (as per the last corrigendum by the European Parliament on 19 April 2024), Especially with the adoption of Art. 18(8)(9) of the EU Directive 2023/2225, of 18 October, which will go into effect on 20 November 2026, there should be more emphasis on the need for hybrid oversight in AI-driven scoring to ensure fairness and transparency. In fact, the range of EU regulations on AI-based consumer credit will soon impact the AI lending industry locally and globally, as shown by the broad territorial scope of AIA's Art. 2. Consequently, engineering the law of consumer's CWA is imperative. Generally, the proposed MAS framework consists of several layers arranged in a specific sequence, as follows: firstly, the Data Layer gathers legitimate predictor sets from traditional sources; then, the Decision Support System Layer, whose Neural Network model is trained using k-fold Cross Validation, provides recommendations based on the feeder data; the eXplainability (XAI) multi-structure comprises Three-Step-Agents; and, lastly, the Oversight Layer has a 'Bottom Stop' for analysts to intervene in a timely manner. From the analysis, one can assure a vital component of this software is the XAY layer. It appears as a transparent curtain covering the AI's decision-making process, enabling comprehension, reflection, and further feasible oversight. Local Interpretable Model-agnostic Explanations (LIME) might act as a pillar by offering counterfactual insights. SHapley Additive exPlanation (SHAP), another agent in the XAI layer, could address potential discrimination issues, identifying the contribution of each feature to the prediction. Alternatively, for thin or no file consumers, the Suggestion Agent can promote financial inclusion. It uses lawful alternative sources such as the share of wallet, among others, to search for more advantageous solutions to incomplete evaluation appraisals based on genetic programming. Overall, this research aspires to bring the concept of Machine-Centered Anthropocentrism to the table of EU policymaking. It acknowledges that, when put into service, credit analysts no longer exert full control over the data-driven entities programmers have given 'birth' to. With similar explanatory agents under supervision, AI itself can become self-accountable, prioritizing human concerns and values. AI decisions should not be vilified inherently. The issue lies in how they are integrated into decision-making and whether they align with non-discrimination principles and transparency rules.