

**THE ANALYTICAL INFORMATION SYSTEM DEVELOPMENT
FOR ASSESSING THE IMPARTIALITY
OF THE RESPONSE OF ARTIFICIAL INTELLIGENCE SYSTEMS**

Sydorenko G.Yu., Gerasymchuk V.D.

National Technical University «Kharkiv Polytechnic Institute», Kharkiv

Artificial intelligence (AI) systems are increasingly being used in the financial sector, particularly to automate credit decision-making. However, such systems can inherit or even reinforce existing social biases, leading to discrimination against certain groups of the population on grounds that should not affect creditworthiness (gender, race, age etc.) [1].

The opacity of algorithms (“black boxes”) makes it difficult to identify and eliminate such biases. Therefore, the development of tools for analyzing and assessing the transparency and impartiality of AI systems in credit scoring is an extremely urgent task, which has both technical and important social significance.

The main goal of the work consists in developing an analytical informational system (AIS), which would allow for comprehensive assessment transparency and identify possible bias in machine learning models that used for credit scoring. The system has use modern metrics and tools in particular functionality libraries AIF 360.

AIF 360 provides a large collection generally accepted fairness metrics (>30) and algorithms mitigation prejudices (>10) that deprives necessity their independent implementation and provides consistency with the best world practices. Use ready, well documented and tested libraries much accelerates process developments analytical systems, allowing focus on AIS architecture and interpretation results [2].

In this article was developed architectural decision analytical informational a system designed to evaluate credit scoring models and create software AIS components responsible for loading data, model training and calculation of fairness metrics received results.

Developed as part of the work analytical information system using libraries AIF 360 is an important step towards ensuring transparency and fairness of AI in a critically important in the field of credit scoring. It provides necessary tools for detection, quantification assessment and potential mitigation discriminatory prejudices, promoting responsible implementation artificial intelligence technologies.

References:

1. Bellamy RK Ket al. AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development* . 2019. Vol. 63, No. 4/5. P. 4:1–4:15. DOI: 10.1147/JRD.2019.2942287.

2. IBM Research. AI Fairness 360. *AIF360 Documentation* . URL: <https://aif360.readthedocs.io/en/stable/>