The Covid pandemic accelerated an already rapidly evolving trend towards a cashless society, and simultaneously led to an even greater demand for more effective online fraud protection measures. The Financial Services industry needs to maintain momentum in fraud prevention by screening online transactions with increasingly sophisticated Machine Learning algorithms. At the same time these service providers do not have a free hand in the implementation of fraud prevention applications, as regulators and public alike still demand accountability from industry AI. This is a theme which looms even larger as the presence of GenAI becomes even more ubiquitous in the modern world, and a certain suspicion around the perceived reliance in AI grows as the first quarter of the 21st century ends.

Credit card fraud detection through ML technology been a commonplace and active area of research for nearly two decades. This is heavily driven by the persistent and mounting threat from ‘bad actors’ in this domain, and the billions of Euros lost each year by individuals and Financial Institutions. Access to better sources of data and more sophisticated models continues to improve detection rates. However, the challenges continues to evolve, and financial institutions increasingly rely on concepts such as Artificial Neural Networks (ANN) to increase the speed and accuracy of credit card fraud detection. Even a cursory review of academic research in the area of CC fraud prevention will shine a light on the dilemma that companies face in this particular domain of crime prevention. It is not enough to stop a case of suspected credit card fraud; a company must be able to explain *why*. Trust in the ‘*black box*’ alone will not suffice. Increasingly, it is not just a case of justifying a decision on fraud to an individual customer, as industry and government regulators will demand that such decision-making processes are transparent and comprehensible.

Companies delivering applications to the financial crime prevention sector have been cognisant in very recent years of the need to add explainability into their product suite. Vendors (such as IBM, Actimize, SymphonyAI, etc.) will boast about advancements in detection rates in areas such as credit card fraud but have also started to supplement these offerings with explanation data for fraud investigators. Auditors will demand that Financial Institutions can demonstrate an evolving process to prevent credit card crime but can also stand over decisions as to why a client’s cards transaction has been delayed/rejected (or why not). That said, commercial Product development of embedded explanations in fraud detection tools is still at a relatively nascent stage.

Data Scientists working in this commercial sphere are aware of the various XAI techniques that can be applied in the domain of credit card fraud, but there is still relatively little published research on the comparative benefits of such approaches in this context. Furthermore, a significant volume of research focuses on the human interpretation of XAI output, whether the subject is fraud detection or health care prediction. Human surveys are costly to implement and can be susceptible to the bias and/or lack of domain knowledge of the participant. The focus of this paper is to look at an automated and statistical comparison of established XAI methods for credit card fraud and assess if a qualitative difference exists between them in terms of general performance.