

Predictive power of early warning system powered by 400k data sources and natural language processing model and its GenAI extension.

#### Abstract

Over the last five years we developed and are continuously enhancing EWS which gathers, aggregates and analyses alternative data sources in near real-time and applies natural language processing techniques to identify potential threats and risks to the credit portfolio. We've built our algorithms to ensure forward-looking prediction of emerging risks to allow the bank time for mitigation. Our Gen AI extension has extended the system benefits drastically, through the use of agent models to refine the output format based on the user's needs and function in the organization.

Since last presenting at CSCCC in 2023, we have improved the methodological robustness as well as the predictive power and added GenAI-driven functionality. In our research and presentation we will focus on our lessons learned as to ways to increase accuracy and transparency of early warning models.

This presentation is the joint work of Deloitte and First Abu Dhabi Bank.

In addition to traditional data sources, our work utilizes news articles, macro indicators and market data. The objective of this research is to identify a set of risk indicators that can be used to predict future credit events using sentiment analysis and artificial intelligence. These risk indicators are sourced from over 400 thousands data sources, and represent millions of articles, that are processed in order to develop the learning algorithm and identify precise events, threats, performance indicators, and causal links based on deep learning techniques and industry experiences.

#### Authors & Affiliations

Dr Natalie Soldatkova<sup>1</sup>, Mr Adnan Mirza<sup>2</sup>, Mr Damian Hales<sup>1</sup>

<sup>1</sup>Deloitte, London, United Kingdom. <sup>2</sup>First Abu Dhabi Bank, Abu Dhabi, UAE