**Project Number:** IPMN-HKU-2024-25-04

**Authorized Institutions:** MOX Bank

**Department:** Technology - (Data, Security and Innovation CDSIO Department)

**Project Title:** Innovating Advanced Methods in Fighting Financial Crime and Credit Risk / Fraud

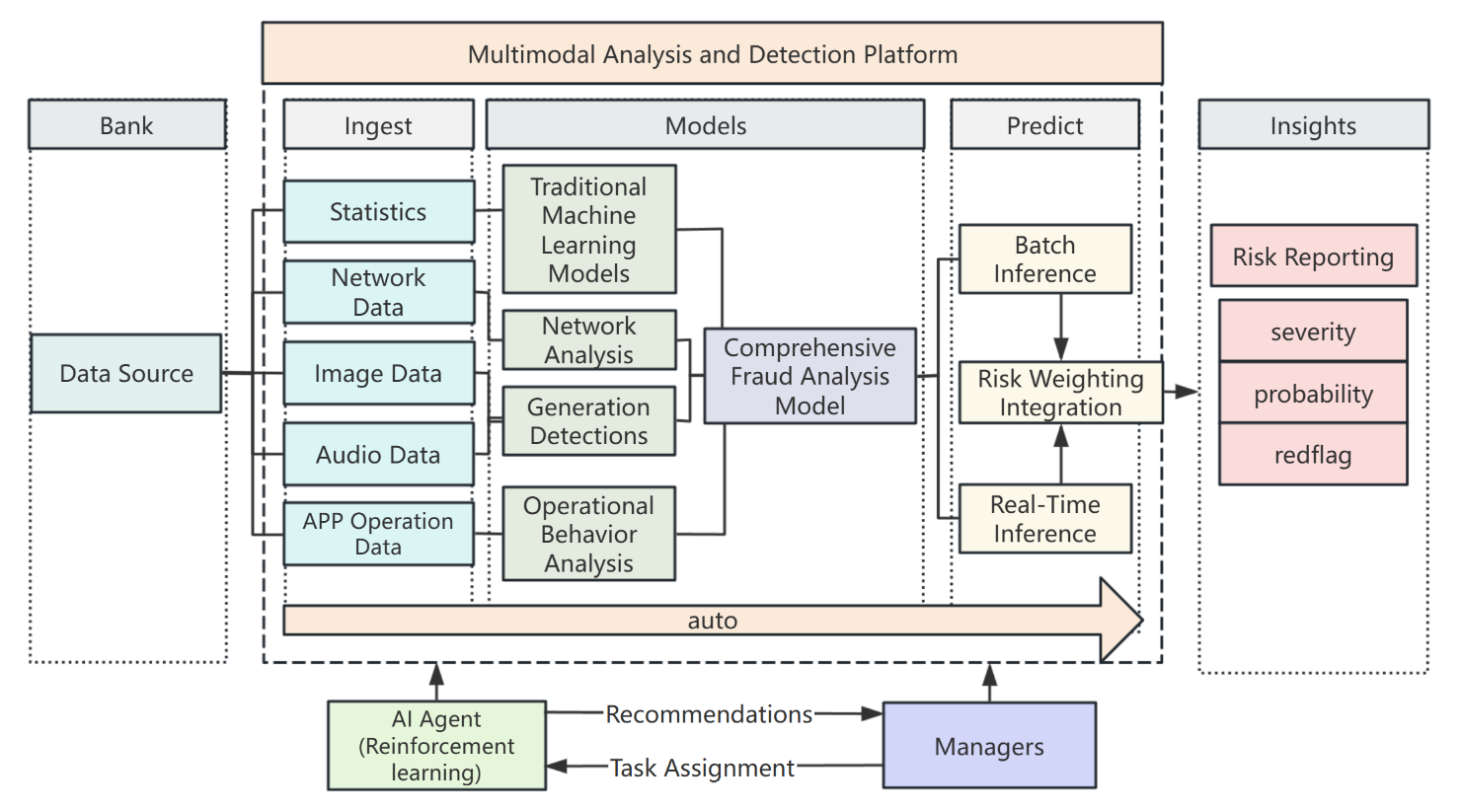
**Group Member:** Deng Dezhao **(Leader)**, Wang Yujin, Yan Chenyue, Yu Caili

**Problem Statement:**

Virtual banks, operating exclusively online, are inherently exposed to higher risks such as cyber threats and data breaches compared to traditional banks with physical branches. Emerging advanced techniques, particularly those leveraging AI and machine learning, have enabled sophisticated fraud methods, posing significant challenges for virtual banks in keeping pace with new threats. Techniques like Stable Diffusion Model generated images are being misused to create convincing deepfake identities and deceive biometric authentication systems, adding another layer of complexity to the risk landscape.

Traditional risk detection approaches, including predictions based solely on a single fixed machine learning model, are insufficient to address these evolving risks. Therefore, there is a critical need for a revolution in risk management strategies, involving the development and implementation of complex models that integrate and analyze multiple data types utilizing innovative AI tools. These models will provide a comprehensive view of potential risks and enable proactive measures to mitigate them effectively, ensuring the security and stability of virtual banking operations.

**Proposed System Architcture:**



We expect to build a Multimodal Analysis and Detection Platform to provide risk reporting insights for Financial Crime and Credit Risk/Fraud.

As shown in the figure, starting with the ingest of data from the bank data sources, we expected multiple types of data will be acquired, including the basic transaction data that is generally used for traditional machine learning models, and network, image, audio data, app operation trace data, etc., which are additionally involved in the comprehensive consideration. Different workflows are taken for different types of data, including the simultaneous use of different models, and then collect the evaluation results into comprehensive models. The administrator can directly control the use of models and the selection of parameters on the platform, or simply specify the task through the AI Agent.

Finally, batch detection of report data can be performed regularly, and risks can also be detected in real-time during the transaction security confirmation process. The output results influence the weights to provide insights to form a final risk report, including potential risk severity and probability, proposing redflag.

**Project Highlights:**

* **Multimodal analysis and detection platform:**

We would use different analysis methods for different data types. We plan to apply various traditional machine learning models to transaction datasets. We would also collect network data and App operation data based on user behavior and use network analysis and operational behavior analysis techniques to assist risk detection. For image and audio data, we will use advanced multimodal models to detect AI-generated contents and increase the sensibility of our platform to fraud cases.

* **AI agent:**

We attach an AI agent based on OS-copilot to facilitate the interaction with our platform. Managers could use natural language to assign tasks and recieve recommendations for specific fraud cases. The platform will adjust the model selection according to the command as needed, and select the parameters to execute the risk detection process

* **Generation detection:**

Processing and analysis of non-traditional data structures (image data/audio data) will be a new source of features for predictive fraud models. At the time of user registration and subsequent transaction execution, the system will detect whether the image/voiceprint verification information provided is AIGC as a feature of fraud.

* **Behavior analysis:**

Based on the idea of funnel analysis. We can identify fraud risks by analyzing the difference between the operational behavior of suspected financial fraudsters and the behavior of normal users. We use indicators such as the order of the content of the app access request, the time interval and frequency of the access, and whether or not the program operates with scripts.

* **Automated detection support:**

Our system platform supports integration into the bank's existing detection process. It can be set up to batch process data on a regular basis to generate risk reports; can detect real-time single-transaction data to provide risk alerts; can handle detection tasks issued by managers as needed.

* **Multidimensional risk report:**

The risk report insights output by our system are multidimensional. Risk prediction scores can be indicated for each feature provide by multimodal models, analyzing the significance of the impact of variables in the model. The processing is transparent and traceable and the results are interpretable. Thresholds can also be set for different parameters to produce graded risk alerts for managers to review and process. Potentially fraudulent transactions and potentially fraudulent users/accounts are also flagged.