Fight against Money Laundering with Optical Character Recognition

Authors : Saikiran Subbagari, Avinash Malladhi

Abstract : Anti Money Laundering (AML) regulations are designed to prevent money laundering and terrorist financing activities worldwide. Financial institutions around the world are legally obligated to identify, assess and mitigate the risks associated with money laundering and report any suspicious transactions to governing authorities. With increasing volumes of data to analyze, financial institutions seek to automate their AML processes. In the rise of financial crimes, optical character recognition (OCR), in combination with machine learning (ML) algorithms, serves as a crucial tool for automating AML processes by extracting the data from documents and identifying suspicious transactions. In this paper, we examine the utilization of OCR for AML and delve into various OCR techniques employed in AML processes. These techniques encompass template-based, feature-based, neural network-based, natural language processing (NLP), hidden markov models (HMMs), conditional random fields (CRFs), binarizations, pattern matching and stroke width transform (SWT). We evaluate each technique, discussing their strengths and constraints. Also, we emphasize on how OCR can improve the accuracy of customer identity verification by comparing the extracted text with the office of foreign assets control (OFAC) watchlist. We will also discuss how OCR helps to overcome language barriers in AML compliance. We also address the implementation challenges that OCR-based AML systems may face and offer recommendations for financial institutions based on the data from previous research studies, which illustrate the effectiveness of OCR-based AML.