Combining Shallow and Deep Unsupervised Machine Learning Techniques to Detect Bad Actors in Complex Datasets

Authors : Jun Ming Moey, Zhiyaun Chen, David Nicholson

Abstract : Bad actors are often hard to detect in data that imprints their behaviour patterns because they are comparatively rare events embedded in non-bad actor data. An unsupervised machine learning framework is applied here to detect bad actors in financial crime datasets that record millions of transactions undertaken by hundreds of actors (<0.01% bad). Specifically, the framework combines 'shallow' (PCA, Isolation Forest) and 'deep' (Autoencoder) methods to detect outlier patterns. Detection performance analysis for both the individual methods and their combination is reported.