# Sequence classification for credit-card fraud detection

## **Abstract**

Due to the growing volume of electronic payments, the monetary strain of credit-card fraud is turning into a substantial challenge for financial institutions and service providers, thus forcing them to continuously improve their fraud detection systems. However, modern data-driven and learning-based methods, despite their popularity in other domains, only slowly find their way into business applications.

In this paper, we phrase the fraud detection problem as a sequence [classification task](https://www.sciencedirect.com/topics/computer-science/classification-task" \o "Learn more about classification task from ScienceDirect's AI-generated Topic Pages) and employ Long Short-Term Memory (LSTM) networks to incorporate transaction sequences. We also integrate state-of-the-art feature aggregation strategies and report our results by means of traditional retrieval metrics.

A comparison to a baseline random forest (RF) classifier showed that the LSTM improves detection accuracy on offline transactions where the card-holder is physically present at a merchant. Both the sequential and non-sequential learning approaches benefit strongly from manual feature aggregation strategies. A subsequent analysis of true positives revealed that both approaches tend to detect different frauds, which suggests a combination of the two. We conclude our study with a discussion on both practical and scientific challenges that remain unsolved.