Abstract

This research focuses on predicting stock returns using the K-Medoids clustering method and the Long Short-Term Memory (LSTM) model. The primary challenge lies in forecasting stock prices, which are then converted into return predictions. Clustering is performed to group stocks with similar price movements, facilitating the preparation of data for training the LSTM model within each cluster. This issue is crucial for aiding investors in making more informed investment decisions by leveraging predictions within specific stock clusters. Through clustering with K-Medoids, based on average returns and return standard deviation, the LSTM model is trained to predict daily returns for each stock within different clusters using the average stock price in each cluster. The data is divided into training (2013-2019) and testing (2020-2022) datasets, with model evaluation conducted using Root Mean Square Error (RMSE). The implementation results indicate prediction performance measured by RMSE for each cluster, with Cluster 3 showing the best performance with a testing RMSE of 0.0300, while Cluster 4 exhibited the worst performance with an RMSE of 0.3995. In the formation of an equal weight portfolio, tested from May 2020 to January 2023, the portfolio value grew from 1 to 2.50, with an average return of 0.0014 and a return standard deviation of 0.0158, indicating potential gains with lower risk compared to the *LQ45* index.

Keywords: K-Medoids, LSTM, *Return* Prediction, portfolio, *Equal Weight*