

ABSTRACT

In this work we made an attempt to predict the upwards/downwards movement of the S&P 500 index from the news articles published by Bloomberg and Reuters. We employed the SVM classifier and conducted multiple experiments aiming at understanding the shape of the data and the specifics of the task better. As a result, we established the common evaluation settings for all our subsequent experiments. After that we tried incorporating various features into the model and also replicated several approaches previously suggested in the literature. We were able to identify some non-trivial dependencies in the data which helped us achieve a high accuracy on the development set. However, none of the models that we built showed comparable performance on the test set. We have come to the conclusion that whereas some trends or patterns can be identified in a particular dataset, such findings are usually barely transferable to other data. The experiments that we conducted support the idea that the stock market is changing at random and a high quality of prediction may only be achieved on particular sets of data and under very special settings, but not for the task of stock market prediction in general.