

File name: FID-013.txt

Result: PLAGIARISM NOT DETECTED

Plagiarism Detected: 0.00%

Text to analyze: Explainable Machine Learning Exploiting News and Domain-Specific Lexicon for Stock Market Forecasting

In this manuscript, we propose a Machine Learning approach to tackle a binary classification problem whose goal is to predict the magnitude (high or low) of future stock price variations for individual companies of the S&P 500 index. Sets of lexicons are generated from globally published articles with the goal of identifying the most impactful words on the market in a specific time interval and within a certain business sector. A feature engineering process is then performed out of the generated lexicons, and the obtained features are fed to a Decision Tree classifier. The predicted label (high or low) represents the underlying company's stock price variation on the next day, being either higher or lower than a certain threshold. The performance evaluation we have carried out through a walk-forward strategy, and against a set of solid baselines, shows that our approach clearly outperforms the competitors. Moreover, the devised Artificial Intelligence (AI) approach is explainable, in the sense that we analyze the white-box behind the classifier and provide a set of explanations on the obtained results.

Sentence: The following sentence: 'Explainable Machine Learning Exploiting News and Domain-Specific Lexicon for Stock Market Forecasting

In this manuscript, we propose a Machine Learning approach to tackle a binary classification problem whose goal is to predict the magnitude (high or low) of future stock price variations for individual companies of the S&P 500 index.' does not present plagiarism

Sentence: The following sentence: 'Sets of lexicons are generated from globally published articles with the goal of identifying the most impactful words on the market in a specific time interval and within a certain business sector.' does not present plagiarism

Sentence: The following sentence: 'A feature engineering process is then performed out of the generated lexicons, and the obtained features are fed to a Decision Tree classifier.' does not present plagiarism

Sentence: The following sentence: 'The predicted label (high or low) represents the underlying company's stock price variation on the next day, being either higher or lower than a certain threshold.' does not present plagiarism

Sentence: The following sentence: 'The performance evaluation we have carried out through a walk-forward strategy, and against a set of solid baselines, shows that our approach clearly outperforms the competitors.' does not present plagiarism

Sentence: The following sentence: 'Moreover, the devised Artificial Intelligence (AI) approach is explainable, in the sense that we analyze the white-box behind the classifier and provide a set of explanations on the obtained results.' does not present plagiarism