



ABSTRACT

The prediction of stock price movement direction is still quite popular among academics because stock values fluctuate and are influenced by various factors, making it difficult to predict. In predicting stock prices, there are several types of analysis used, namely technical analysis using technical indicators (TI), fundamental analysis using financial factors and the economic condition of the company concerned, and sentiment analysis using sentiment taken from various sources such as social media, news, and others. Technical analysis is the analysis chosen in this study, considering that fundamental analysis and sentiment analysis do not have stable data.

Long -Term Memory (LSTM) is one of the most popular methods used to forecast stocks; this happens because LSTM is believed to be proficient in processing time series. However, LSTM has weaknesses, namely frequent overfitting and high computational complexity because there are many hyperparameters that must be learned by the user. The overfitting problem occurs because the data dimension contains too much irrelevant information for the training model to process. The use of irrelevant IT is one of the causes of overfitting. From these problems, pre-processing is needed to reduce the dimensions of the attributes used using the Principal Component Analysis (PCA) method. The computational complexity related to the number of hyperparameters can be overcome by using hyperparameter tuning, which allows the model to match the right hyperparameters so as to optimize the accuracy of the results. The hyperparameters that will be combined using hyperparameter tuning are the number of layers, batch size, neurons, learning rate, and dropout.

The goal of this study was to identify the categorization model that would be most effective in predicting the movement of 20 stocks, separated into blue-chip stocks and small-cap companies, in one direction or another. In contrast to utilizing a dataset TI, it was discovered that using a dataset as an input produced superior average accuracy. The use of dataset gives an average accuracy optimization to the LSTM architecture layer, and 0.25% 4-layer LSTM architecture. Then the use of



hyperparameter tuning was able to produce an average accuracy optimization of 3.19% for the 2-layer LSTM architecture model , and 0.28% for the model with hyperparameter tuning and 4-layer LSTM architecture . In addition, the use of Hyperparameter Tuning to find the best combination of hyperparameters for each stock is able to have an increasing effect on an average of 75% of stocks with an average maximum increase accuracy of 35.17%.

Keyword – LSTM, *Stocks Price*, *Technical Indicators*, PCA, *Tuning*, *Technical Analysis*



INTISARI

Prediksi pergerakan arah harga saham masih menjadi hal yang cukup populer di kalangan akademisi karena nilai saham yang bergerak secara fluktuatif dan dipengaruhi oleh berbagai aspek sehingga sulit diprediksi. Dalam memprediksi harga saham ada beberapa jenis analisis yang digunakan yaitu analisis teknikal yang menggunakan *Technical Indicators* (TI), analisis Fundamental menggunakan faktor keuangan dan kondisi ekonomi perusahaan terkait, analisis sentimen menggunakan sentimen yang diambil dari berbagai sumber seperti media sosial, berita dan lain-lain. Analisis teknikal merupakan analisis yang dipilih dalam penelitian ini mempertimbangkan bahwa analisis fundamental dan analisis sentimen tidak memiliki data yang stabil.

Long Short-Term Memory (LSTM) menjadi salah satu metode yang paling populer digunakan untuk melakukan peramalan saham, hal ini terjadi karena LSTM yang dipercaya mahir dalam memproses deret waktu. Namun LSTM memiliki kelemahan yaitu sering terjadi *overfitting* dan kompleksitas komputasi yang tinggi karena banyak hyperparameter yang harus dipelajari oleh pengguna. Masalah *overfitting* terjadi karena dimensi data terlalu banyak berisi informasi tidak relevan untuk diproses oleh model pelatihan. Penggunaan TI yang tidak relevan merupakan salah satu penyebab terjadinya *overfitting*. Dari masalah tersebut kemudian dibutuhkan *pra-processing* untuk mengurangi dimensi dari atribut yang digunakan menggunakan metode *Principal Component Analysis* (PCA). Kompleksitas komputasi terkait banyaknya hyperparameter dapat diatasi dengan menggunakan *hyperparameter tuning* yang memungkinkan model untuk mencocokkan hyperparameter yang tepat sehingga mengoptimalkan hasil akurasi. Hyperparameter yang akan dikombinasikan menggunakan *hyperparameter tuning* adalah jumlah *layer*, *batch size*, *neuron*, *learning rate* dan *dropout*.

Penelitian ini dilakukan untuk mencari model klasifikasi yang paling tepat untuk prediksi arah pergerakan 20 saham yang dibedakan menjadi saham *blue chip*



dan saham *small cap*. Ditemukan bahwa penggunaan *dataset* PCA sebagai masukan menghasilkan rata-rata akurasi yang lebih baik daripada menggunakan masukan *dataset* TI. Penggunaan *dataset* PCA memberi optimasi rata-rata akurasi sebesar 0,15% untuk model dengan hyperparameter default dan arsitektur LSTM 2 *layer*, dan 0,25% arsitektur LSTM 4 *layer*. Kemudian untuk penggunaan *hyperparameter tuning* mampu menghasilkan optimasi akurasi rata-rata sebesar 3,19% untuk model arsitektur LSTM 2 *layers*, dan 0,28% untuk model dengan *hyperparameter tuning* dan LSTM arsitektur 4 *layers*. Selain itu, penggunaan *Hyperparameter Tuning* untuk menemukan kombinasi hyperparameter terbaik untuk setiap saham mampu memberi pengaruh kenaikan terhadap rata-rata 75% saham dengan rata-rata akurasi kenaikan maksimal sebesar 35,17%.

Kata kunci – LSTM, Saham, *Technical Indicators*, PCA, Klasifikasi