

ABSTRAK

Kegiatan investasi sangatlah umum dan diminati masyarakat untuk memperoleh keuntungan dari sejumlah dividen atau kenaikan harga saham di masa yang akan datang. Namun, investor tidak akan selalu mengalami keuntungan, karena kegiatan investasi memiliki unsur ketidakpastian dan berbagai risiko. Analisis dan pertimbangan pun perlu dilakukan sebelum berinvestasi pada emiten tertentu, untuk mengetahui apakah besar keuntungan, risiko, dan kemampuan berinvestasi yang dimiliki untuk menanggung risiko adalah sepadan. Risiko dapat diukur dengan *Value at Risk* (VaR) untuk mengetahui kerugian maksimum yang mungkin terjadi. Keuntungan dapat diukur dengan *expected return*, dan volatilitas saham dapat digunakan untuk mengestimasi risiko serta keuntungan yang akan diperoleh dalam kurun waktu tertentu. Kemudian, *K-Means Clustering* dapat digunakan untuk menentukan kelompok emiten pada indeks LQ 45 yang layak diinvestasikan, berdasarkan *expected return*, volatilitas saham, dan *Value at Risk* (VaR). Namun, terdapat korelasi yang sangat kuat antara volatilitas saham dan *Value at Risk* (VaR). Sehingga, selain *Euclidean Distance* untuk perhitungan jarak pada *K-Means*, penelitian ini akan melibatkan *Mahalanobis Distance* dan *Principal Component Analysis* (PCA) yang dapat mengatasi adanya korelasi antar variabel. Penelitian ini ditujukan untuk menentukan jumlah *cluster*, mengetahui karakteristik dari *cluster* yang terbentuk, serta membandingkan performa *K-Means Clustering* saat perhitungan jarak menggunakan metode *Euclidean Distance*, *Mahalanobis Distance*, dan menerapkan *Principal Component Analysis* (PCA) sebelum *clustering*. Hasilnya, jumlah *cluster* yang optimal adalah sebanyak tiga, karakteristik antar *clusters* yang diperoleh kurang berbeda atau kurang heterogen, dan performa *K-Means Clustering* dengan penerapan PCA lebih baik dibandingkan dengan *Euclidean Distance* saja dan *Mahalanobis Distance*. Namun, *K-Means Clustering* dengan *Euclidean Distance* saja lebih baik dibandingkan dengan *Mahalanobis Distance*.

Kata kunci: *K-Means Clustering*, *Elbow Method*, *Euclidean Distance*, *Mahalanobis Distance*, *Principal Component Analysis* (PCA), *Expected Return*, Volatilitas Saham, *Value at Risk* (VaR) dengan *Historical Simulation Method*

ABSTRACT

Investment activities are very common and in demand by the public to acquire benefit from dividends or an increase in stock prices in the future. However, investors will not always experience profits, because investment activities have elements of uncertainty and various risks. Analysis and considerations also need to be carried out before investing in a particular issuer, to find out whether the amount of profit, risk, and investment ability that is owned are worth and able to bear the risk. Risk can be measured by Value at Risk (VaR) to determine the maximum possible loss. Profits can be measured by the expected return, and stock volatility can be used to estimate the risks and benefits that will be obtained within a certain period of time. Then, K-Means Clustering can be used to determine the group of issuers on the LQ 45 index that are worth investing in, based on expected returns, stock volatility, and Value at Risk (VaR). However, there is a very strong correlation between stock volatility and Value at Risk (VaR). So, beside Euclidean Distance to calculate distance in K-Means, this research will involve Mahalanobis Distance and Principal Component Analysis (PCA) which can overcome the correlation between variables. This study aims to determine the number of clusters, determine the characteristics of the clusters formed, and compare the performance of K-Means Clustering when calculating distances using the Euclidean Distance, Mahalanobis Distance, and applying Principal Component Analysis (PCA) before clustering. As a result, the optimal number of clusters is three, the characteristics between the clusters obtained are less different or less heterogeneous, and the performance of K-Means Clustering with PCA application is better than Euclidean Distance alone and Mahalanobis Distance. However, K-Means Clustering with Euclidean Distance alone is better than Mahalanobis Distance.

Keywords: K-Means Clustering, Elbow Method, Euclidean Distance, Mahalanobis Distance, Principal Component Analysis (PCA), Expected Return, Volatilitas Saham, Value at Risk (VaR) with Historical Simulation Method