

## INTISARI

### RANCANG BANGUN *DASHBOARD* SISTEM INFORMASI DAN PREDIKSI PRODUKSI TANAMAN PANGAN BERBASIS *WEB* MENGGUNAKAN ALGORITMA KMEANS DAN *LINEAR REGRESSION*

Ivon Virginia Patala

20/464394/SV/18713

Dalam Pilar Pembangunan Indonesia 2045, ketahanan pangan merupakan salah satu isu utama yang perlu diperhatikan proses pemberdayaannya. Sektor pertanian sangat berperan dalam memenuhi kebutuhan pangan yang merupakan kebutuhan primer. Menurut data Badan Pusat Statistik, produksi tanaman pangan di Indonesia dalam rentang tahun 1993 hingga 2023 menunjukkan perubahan luas panen setiap tahunnya yang mempengaruhi hasil produksi. Pemerintah akan terus berupaya untuk meningkatkan produksi tanaman pangan. Untuk mendukung hal tersebut, dibutuhkan sebuah sistem dengan tujuan memantau performa produksi menggunakan teknik klasterisasi dan regresi. Data yang akan digunakan dalam proses analisis meliputi tujuh komoditas tanaman pangan yaitu padi, jagung, kedelai, kacang tanah, kacang hijau, ubi jalar, dan ubi kayu yang setiap komoditasnya berjumlah 986 data. Teknik klasterisasi yaitu mengelompokkan data berdasarkan ciri-ciri yang dimilikinya menggunakan algoritma KMeans. Klaster ditentukan berdasarkan luas panen dan produksi yang dibagi menjadi tiga klaster yaitu rendah, sedang, dan tinggi. Berdasarkan perbandingan performa, algoritma KMeans lebih optimal dibandingkan dengan algoritma KMedoids dibuktikan dengan nilai matriks evaluasi *Silhouette Coefficient* mendekati 1 dengan rentang nilai 0,69 hingga 0,87, nilai *Davies Bouldin Index* yang rendah yaitu dalam rentang nilai 0,33 hingga 0,49, dan nilai *Calinski Harabasz Index* yang lebih tinggi dengan rentang nilai 555,91 hingga 1889,96 pada setiap komoditas. Pada Analisis menggunakan algoritma *Linear Regression*, nilai luas panen memiliki hubungan yang linier dengan nilai produksi dibuktikan dengan nilai *Coefficient of Determination* ( $R^2$ ) yang mendekati nilai 1 yaitu dalam rentang nilai 0,84 hingga 0,99 untuk komoditas padi, jagung, kedelai, kacang tanah, kacang hijau, dan ubi kayu, sedangkan pada komoditas ubi jalar memiliki hubungan yang rendah yaitu dengan nilai 0,37. Nilai luas panen juga berhasil digunakan untuk memprediksi nilai produksi dengan akurasi yang baik dibuktikan dengan nilai *Root Mean Square Error* (RMSE) pada rentang 6320,13 hingga 408837,43 untuk setiap komoditas. Hasil analisis akan ditampilkan dalam bentuk dashboard visualisasi menggunakan bahasa pemrograman *python* yaitu memanfaatkan *library streamlit*. Sistem ini diharapkan dapat membantu pemerintah dalam memantau produksi tanaman pangan di Indonesia dan mengambil keputusan untuk mengatasi permasalahan yang ada.

Kata kunci: Tanaman Pangan, Klasterisasi, KMeans, Regresi, *Linear Regression*

## ABSTRACT

### **DESIGN AND DEVELOPMENT OF A WEB-BASED CROP PRODUCTION PREDICTION AND INFORMATION SYSTEM DASHBOARD USING KMEANS AND LINEAR REGRESSION ALGORITHMS**

Ivon Virginia Patala

20/464394/SV/18713

*In Indonesia's Development Pillar 2045, food security is one of the main issues that needs to be considered in the empowerment process. The agricultural sector plays a major role in meeting food needs, which are primary needs. According to data from the Central Bureau of Statistics, food crop production in Indonesia from 1993 to 2023 shows changes in harvest areas every year which affects production yields. The government will continue to strive to increase food crop production. To support this, a system is needed with the aim of monitoring production performance using clustering and regression techniques. The data that will be used in the analysis process includes seven food crop commodities, namely rice, corn, soybeans, peanuts, green beans, sweet potatoes, and cassava, each commodity totaling 986 data. The clustering technique is to group the data based on its characteristics using the KMeans algorithm. Clusters are determined based on harvest area and production which are divided into three clusters, namely low, medium and high. Based on the performance comparison, the KMeans algorithm is more optimal than the KMedoids algorithm as evidenced by the Silhouette Coefficient evaluation metrics value close to 1 with a value range of 0.69 to 0.87, a low Davies Bouldin Index value in the value range of 0.33 to 0.49, and a higher Calinski Harabasz Index value with a value range of 555.91 to 1889.96 for each commodity. In the analysis using the Linear Regression algorithm, the value of the harvest area has a linear relationship with the value of production as evidenced by the value of the Coefficient of Determination ( $R^2$ ) which is close to the value of 1, namely in the range of values 0.84 to 0.99 for rice, corn, soybeans, peanuts, green beans, and cassava commodities, while the sweet potato commodity has a low relationship with a value of 0.37. The value of the harvest area was also successfully used to predict the value of production with good accuracy as evidenced by the Root Mean Square Error (RMSE) value in the range of 6320.13 to 408837.43 for each commodity. The analysis results will be displayed in the form of a visualization dashboard using the python programming language, utilizing the streamlit library. This system is expected to help the government in monitoring food crop production in Indonesia and making decisions to overcome existing problems.*

**Keywords:** Food Crops, Clustering, KMeans, Regression, Linear Regression