

ABSTRAK

Tugas akhir ini dilatar belakangi oleh kebutuhan dalam pemodelan regresi untuk menemukan model terbaik dengan waktu yang singkat dengan jumlah data yang relatif besar. Metode validasi silang merupakan metode yang efektif untuk mencari model terbaik dengan jumlah data yang besar sekaligus untuk mengevaluasi model. Secara lebih detail pada skripsi ini akan digunakan validasi silang K-lipat untuk menentukan model terbaik berdasarkan data Badan Pusat Statistik (BPS) seluruh provinsi di Indonesia pada tahun 2021 mengenai sektor pertanian padi dengan pendapatan daerah sebagai variabel terikat dan produktifitas serta luas lahan sebagai variabel bebas untuk membentuk model regresi berganda. Kemudian dilakukan metode validasi silang untuk mencari model terbaik serta mengevaluasi model dengan $K = 5$ dan $K = 10$. Berdasarkan hasil evaluasi tersebut akan dilihat bagaimana hasil evaluasi dan model terbaik dalam pendugaan model regresinya.

Kata kunci: *analisis regresi, metode validasi silang, metode validasi silang K-lipat*

ABSTRACT

This final project aims to tackle the challenge of finding the best model for regression modeling in a short time, while dealing with a relatively large amount of data. To achieve this goal, the thesis will employ the cross-validation method, which is an effective technique for both identifying the best model and evaluating its performance using large datasets. Specifically, this project will use the k-fold cross-validation approach to develop a multiple regression model based on Central Statistics Agency (BPS) data from all provinces in Indonesia in 2021. The model will focus on the rice farming sector, using regional income as the dependent variable and productivity and land area as the independent variables. Next, the model will be evaluated using K=5 and K=10 cross-validation methods to determine the best model and assess its performance. Based on the evaluation results, we will be able to identify the best model that accurately estimates the regression model.

Keywords: *regression analysis, cross validation, k-fold cross validation.*