

ABSTRAK

Nama : Ferdiansyah
NIM : 41519010187
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementasi Data Mining Untuk Memprediksi Penjualan Benih Padi Pada PT. XYZ Menggunakan Algoritma C4.5 dan Regresi Linier
Pembimbing : Eliyani, Dr. Ir.

Prediksi penjualan merupakan salah satu strategi untuk melihat peluang penjualan yang ada, yang diharapkan dapat meningkatkan pendapatan suatu perusahaan. Sektor penjualan benih padi merupakan salah satu sektor yang paling dominan bagi Indonesia di karenakan Indonesia negara agraris. Penelitian ini akan berfokus pada penerapan algoritma C4.5 dan Linier Regresi dalam memprediksi hasil penjualan benih padi. Dataset yang digunakan merupakan data laporan penjualan setiap kantor wilayah yang tersebar di seluruh Indonesia dari tahun 2020-2023, yang berasal dari perusahaan PT.XYZ. Lalu akan dilakukan beberapa metode Pre-processing, agar data yang digunakan dapat di evaluasi untuk melakukan proses implementasi algoritma C4.5 dan Regresi Linier. Evaluasi model pengujian menggunakan Mean Squared Error (MSE), Mean Absolute Error (MAE), untuk membandingkan hasil nilai pengujian anatara kedua algoritma mana yang baik untuk digunakan pada penelitian ini. Berdasarkan hasil skenario eksperimen yang sudah di lakukan menunjukan bahwa algoritma C4.5 lebih baik dibanding dengan algoritma Regresi Linier dengan besaran data train 70 dan data test 30, mendapatkan hasil pengujian pada data uji nilai *R-Squared* sebesar 0.16, nilai MAE sebesar 51980539.69, dan nilai MSE sebesar 2737595355333400.00. Kemudian hasil uji eksperimen algoritma Regresi Linier mendapatkan hasil *R-Squared* sebesar 0.11, MAE sebesar 66451074.39, dan MSE sebesar 29245676909442588.00.

Kata Kunci:

Data Mining, Regresi Linier, C4.5, Prediksi, Penjualan Benih Padi

ABSTRACT

Name : Ferdiansyah
NIM : 41519010187
Study Program : Informatics Engineering
Title Thesis : Implementation of Data Mining to Predict Rice Seed Sales at PT. XYZes Using C4.5 Algorithm and Linier Regression
Counsellor : Eliyani, Dr. Ir.

Sales prediction is one of the strategies to see existing sales opportunities, which are expected to increase the revenue of a company. The rice seed sales sector is one of the most dominant sectors for Indonesia because Indonesia is an agricultural country. This research will focus on the application of the C4.5 algorithm and Liner Regression in predicting the sales results of rice seeds. The dataset used is the sales report data of each regional office spread throughout Indonesia from 2020-2023, which comes from the company PT.XYZ. Then several pre-processing methods will be carried out, so that the data used can be evaluated to carry out the implementation process of the C4.5 algorithm and Linear Regression. Evaluation of the test model using Mean Squared Error (MSE), Mean Absolute Error (MAE), to compare the results of the test values between the two algorithms which are good for use in this study. Based on the results of experimental scenarios that have been carried out, it shows that the C4.5 algorithm is better than the Linear Regression algorithm with the amount of train data 70 and test data 30, getting test results on test data R-Squared value of 0.16, MAE value of 51980539.69, and MSE value of 27375953555333400.00. Then the results of the Linear Regression algorithm experiment test obtained an R-Squared result of 0.11, MAE of 66451074.39, and MSE of 29245676909442588.00.

Keywords:

Data Mining, Regression Linear, C4.5, Predict, sales of rice seeds