

ABSTRAK

Nama : Denise Paulina Yunita
NIM : 41519010090
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Penerapan Data Mining Dalam Memprediksi
Penyebaran Kasus Covid-19 di Wilayah Kelapa Gading
Menggunakan Algoritma C 4.5
Pembimbing : Rahmat Rian Hidayat, ST., MMSI

Penelitian ini membahas tentang penerapan data mining dalam memprediksi penyebaran COVID-19 di wilayah Kelapa Gading, Jakarta Utara, dengan tujuan untuk membantu mengendalikan kasus COVID-19 di wilayah tersebut. Menurut WHO, COVID-19 merupakan suatu penyakit menular yang disebabkan oleh suatu varian baru dari coronavirus. Penyebaran yang tidak terkendali dapat mengganggu aktivitas sosial dan ekonomi masyarakat. Sementara itu, pelayanan kesehatan dan sumber daya manusia yang terlibat dalam penanganan pandemi COVID-19 di Indonesia masih belum memadai, sementara kasus terus meningkat. Metode yang digunakan adalah algoritma C4.5 dan pengujian validasi menggunakan K Fold sebanyak 10 fold. Hasil penelitian menunjukkan bahwa model menggunakan algoritma C4.5 memiliki tingkat akurasi yang baik, mencapai 0,98. Validasi dengan K Fold menghasilkan akurasi sebesar 0,99 dan hasil pengujian data juga sesuai dengan prediksi model. Berdasarkan hasil prediksi, jika jumlah kasus positif harian kurang dari 35,50, kelasnya adalah "rendah". Jika jumlah kasus positif harian antara 35,50 hingga 50,50, kelasnya adalah "sedang". Dan jika jumlah kasus positif harian lebih dari 50,50, kelasnya adalah "tinggi".

Kata Kunci: Algoritma C4.5, COVID-19, K Fold Validation

ABSTRACT

Name : Denise Paulina Yunita
NIM : 41519010090
Study Program : Informatics Engineering
Title Thesis : Application of Data Mining in Predicting the Spread of Covid-19 Cases in the Kelapa Gading Region Using the C 4.5 Algorithm
Counsellor : Rahmat Rian Hidayat, ST., MMSI

This study explores the utilization of data mining for predicting the spread of COVID-19 in the Kelapa Gading region, North Jakarta, with the aim of helping to control COVID-19 cases in that area. According to the WHO, According to WHO, COVID-19 is an infectious disease caused by a new variant of coronavirus. Uncontrolled spread can disrupt social and economic activities. Meanwhile, healthcare services and the human resources involved in handling the COVID-19 pandemic in Indonesia are still insufficient, while cases continue to increase. The method used is the C4.5 algorithm, and the validation testing is done using K Fold with 10 folds. The research findings show that the model using the C4.5 algorithm has a good level of accuracy, reaching 0.98. Validation with K Fold resulted in an accuracy of 0.99, and the testing data also aligned with the model's predictions. Based on the prediction results, if the daily number of positive cases is less than 35.50, the class is classified as "low". If the daily number of positive cases is between 35.50 and 50.50, the class is classified as "medium". And if the daily number of positive cases is above 50.50, the class is classified as "high".

Keywords: Algorithm C4.5, COVID-19, K Fold Validation