

**PENGARUH TEKNIK *SMOTE* TERHADAP PREDIKSI HARAPAN
HIDUP PENDERITA PENYAKIT HEPATITIS MENGGUNAKAN
ALGORITMA *K- NEAREST NEIGHBOR***

Devi Putri Prianggi¹, Agung Nilogiri², Reni Umilasari³

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah

Jember

E-Mail : deviputri382@gmail.com¹, agungnilogiri@unmuhjember.ac.id²,

reni.umilasari@unmuhjember.ac.id³

ABSTRAK

Hepatitis merupakan penyakit yang menyerang organ hati pada manusia, biasanya disebabkan oleh infeksi jamur, bakteri, penggunaan alkohol, *autoimmune*, dan obat-obatan. Hepatitis sendiri terbagi menjadi 5, yaitu hepatitis A, B, C, D, dan E. Hepatitis B memiliki prevalensi tertinggi kedua di Asia Tenggara. Tujuan dari penelitian ini yaitu untuk mengetahui tingkat akurasi, presisi, dan recall. Dataset yang digunakan berasal dari *Kaggle* berjumlah 142 dan atribut sebanyak 20. Pada uji coba penelitian ini menggunakan *Cross Fold Validation* dengan nilai $K = 2, 4,$ dan 5. Dari implementasi menggunakan algoritma *K-Nearest Neighbor* tanpa *SMOTE* diperoleh hasil validasi terbaik dengan akurasi tertinggi didapatkan 89,66% pada Kfold 5, presisi 78,12% pada Kfold 4, dan recall 94,44% pada Kfold 5. Sedangkan implementasi yang dilakukan menggunakan algoritma *K-Nearest Neighbor* menggunakan *SMOTE* diperoleh hasil validasi terbaik dengan akurasi tertinggi didapatkan 93,48 % pada Kfold 5, presisi 95,00 % pada Kfold 5, dan recall 92,11 % pada Kfold 5.

Kata Kunci : Prediksi Harapan Hidup, *K- Nearest Neighbor*, Teknik *SMOTE*, Hepatitis, *Cross Validation*

**PENGARUH TEKNIK SMOTE TERHADAP PREDIKSI HARAPAN
HIDUP PENDERITA PENYAKIT HEPATITIS MENGGUNAKAN
ALGORITMA K- NEAREST NEIGHBOR**

Devi Putri Prianggi¹, Agung Nilogiri², Reni Umilasari³

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah
Jember

E-Mail : deviputri382@gmail.com¹, agungnilogiri@unmuhjember.ac.id²,
reni.umilasari@unmuhjember.ac.id³

ABSTRACT

Hepatitis is a disease that attacks the liver in humans, usually caused by fungal infections, bacteria, alcohol use, autoimmune, and drugs. Hepatitis itself is divided into 5, namely hepatitis A, B, C, D, and E. Hepatitis B has the second highest prevalence in Southeast Asia. The purpose of this study is to determine the level of accuracy, precision, and recall. The dataset used comes from Kaggle totaling 142 and 20 attributes. In this research trial using Cross Fold Validation with values of $K = 2, 4, \text{ and } 5$. From the implementation using the K-Nearest Neighbor algorithm without SMOTE, The best validation results were obtained with the highest accuracy obtained 89.66% on Kfold 5, 78.12% precision on Kfold 4, and 94.44% recall on Kfold 5. While the implementation carried out using the K-Nearest Neighbor algorithm using SMOTE obtained validation results The best with the highest accuracy obtained 93.48% on Kfold 5, 95.00% precision on Kfold 5, and recall 92.11% on Kfold 5.

Keywords: *Life Expectancy Prediction, K- Nearest Neighbor, SMOTE Technique, Hepatitis, Cross Validation*