KLASIFIKASI *RANDOM FOREST* DENGAN PERBAIKAN *MISSING VALUE* *K-NEAREST NEIGHBOR IMPUTATION* (KNNI) PADA PENYAKIT DIABETES

Skripsi

Untuk Memenuhi Persyaratan

Dalam Menyelesaikan Strata-1 Ilmu Komputer

Oleh

SARTIKA DEWI

1711016120019

PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU

MEI 2023

KLASIFIKASI *RANDOM FOREST* DENGAN PERBAIKAN *MISSING VALUE* *K-NEAREST NEIGHBOR IMPUTATION* (KNNI) PADA PENYAKIT DIABETES

Skripsi

Untuk Memenuhi Persyaratan

Dalam Menyelesaikan Strata-1 Ilmu Komputer

Oleh

SARTIKA DEWI

1711016120019

PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU

MEI 2023

SKRIPSI

KLASFIKASI *RANDOM FOREST* DENGAN PERBAIKAN *MISSING VALUE K-NEAREST NEIGHBOR IMPUTATION* (KNNI) PADA PENYEKIT DIABETES

Oleh:

SARTIKA DEWI

NIM 1711016120019

Telah dipertahankan di depan Dosen Penguji pada Tanggal 23 Mei 2023

Susunan Dosen Penguji:

Pembimbing I	Dosen Penguji I
Triando Hamonangan Saragih, S.Kom., M.Kom	**Dodon Turianto Nugrahadi, S.Kom., M.Eng**
NIP. 199308242019031012	**NIP. 198001122009121002**
Pembimbing II	Dosen Penguji II
Radityo Adi Nugroho, S.T., M.Kom	**Dwi Kartini, S.Kom., M.Kom.**
NIP. 198212042008011006	**NIP. 198704212012122003**

Banjarbaru, 23 Mei 2023

Koordinator Program Studi Ilmu Komputer

Irwan Budiman, S.T., M.Kom,

NIP. 197703252008121001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar keserjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 23 Mei 2023

Yang Menyatakan,

Sartika Dewi
NIM. 1711016120019

ABSTRAK

KLASIFIKASI *RANDOM FOREST* DENGAN PERBAIKAN *MISSING VALUE K-NEAREST NEIGHBOR IMPUTATION* (KNNI) PADA PENYAKIT DIABETES

(Oleh: Sartika Dewi; Pembimbing: Triando Hamonangan Saragih, S.Kom., M.Kom dan Radtyo Adi Nugroho, S.T., M.Kom; 2022; 67 halaman)

Pada saat melakukan survei data penghilangan nilai atau tidak lengkapnya data dapat terjadi, sehingga metode analisis yang tersedia hanya dapat bekerja dengan data lengkap. Pada penelitian ini melakukan penghilangan nilai dengan jumlah proporsi penghilangan yaitu 10%, 20% dan 30%. Tujuan dari penelitian ini yaitu untuk mengetahui hasil akurasi dari imputasi data hilang dengan K-NNI menggunakan klasifikasi *Random Forest* dan klasifikasi *Random Forest* tanpa imputasi. Metode penelitian dilakukan dengan Dataset publik *Pima Indian Diabetes* lalu dengan data yang lengkap dilakukan pengosongan data menggunakan MCAR *(Missing Completely At Random).* Kemudian, dilakukan 2 model yaitu yang pertama imputasi data hilang dengan K-NNI dan klasifikasikan dengan *Random Forest*. Kedua, klasifikasi *Random Forest* tanpa imputasi. Berdasarkan dari hasil akurasi K-NNI menggunakan klasifikasi *Random Forest* didapatkan akurasi terbaik pada penghilangan nilai sebanyak 10% dan nilai K = 6 yaitu sebesar 82,17% dengan presisi 76,71%, *recall* 70,00% dan *f1-Score 73,20%*. Kemudiann untuk hasil akurasi dari klasifikasi *Random Forest* tanpa imputasi didapatkan pada penghilangan nilai 10% sebesar 80,00% dengan presisi 64,52%, *recall* 68,97% dan *f1-Score* 66,67%. Namun, pada pengujian missing data tanpa imputasi K-NNI menggunakan klasifikasi *Random Forest* memperoleh hasil yang tidak lebih baik dari hasil akurasi K-NNI menggunakan klasifikasi *Random Forest.*

Kata Kunci: *Pima Indian Diabetes, missing value, K-Nearest Neighbor Imputation, Random Forest.*

ABSTRACT

CLASSIFICATION OF RANDOM FORESTS WITH MISSING VALUE IMPUTATION OF K-NEAREST NEIGHBOR (KNNI) IN DIABETES DIABETES (*By:* Sartika Dewi; *Supervisor:* Triando Hamonangan Saragih, S.Kom., M.Kom *and* Radtyo Adi Nugroho, S.T., M.Kom; 2022; 67 *pages*)

When conducting a data survey, omissions or incomplete data may occur, so that the available analytical methods can only work with complete data. In this study, devaluation was carried out with the proportion of omissions namely 10%, 20% and 30%. The purpose of this research is to find out the results of the accuracy of imputation of missing data with K-NNI using Random Forest classification and Random Forest classification without imputation. The research method was carried out using the Pima Indian Diabetes public dataset, and then with complete data, blank data was carried out using MCAR (Missing Completely At Random). Then, 2 models were carried out, namely the first was imputed missing data with K-NNI and classified with Random Forest. Second, the classification of Random Forest without imputation. Based on the results of the K-NNI accuracy using the Random Forest classification, the best accuracy was obtained for 10% omission and a K = 6 value of 82.17% with a precision of 76.71%, a recall of 70.00% and an f1-Score of 73.20 %. Then for the results of the accuracy of the Random Forest classification without imputation obtained at 10% omission of 80.00% with a precision of 64.52%, a recall of 68.97% and an f1-Score of 66.67%. However, the missing data test without K-NNI imputation using the Random Forest classification obtained results that were no better than the K-NNI accuracy results using the Random Forest classification.

Keywords*: Pima Indian Diabetes, missing value, K-Nearest Neighbor Imputation, Random Forest.*

PRAKATA

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul "**KLASIFIKASI *RANDOM FOREST* DENGAN PERBAIKAN *MISSING VALUE K-NEAREST NEIGHBOR IMPUTATION* (KNNI) PADA PENYAKIT DIABETES** " untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa pula penulis panjatkan shalawat dan salam ke hadirat Rasulullah Muhammad SAW beserta para sahabat, keluarga, dan pengikut beliau hingga *yaumul qiyamah.*

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut :

1. Keluarga terutama kepada Orang tua yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian skripsi ini.
2. Bapak Triando Hamonangan Saragih, S.Kom,. M.Kom selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran penyelesaian skripsi ini.
3. Bapak Radityo Adi Nugroho, S.T., M.Kom selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Bapak Irwan Budiman, S,T., M.Kom selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
5. Seluruh Dosen dan staf Program Studi Ilmu Komputer FMIPA UNLAM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
6. Teman-teman keluarga Ilmu Komputer angkatan 2017, terima kasih untuk canda, tawa, perjuangan yang sudah dilewati bersama, untuk semua kenangan manis yang telah terukir selama ini. Senang bisa menjadi salah satu bagian dari kehidupan kalian.

7. Terima kasih juga kepada teman-teman yang sudah mendukung saya dalam menyelesaikan penelitian ini, terima kasih kepada Noor Hidayah yang sudah banayk membantu dalam segala hal, terim kasih juga kepada Syaoki Paradisa yang sudah membantu saya dalam proses penelitian saya, terima kasih banyak juga kepada Sulastri Nor Indah Sari, Mia Lidiyani, Regina, dan Nurhuda dan Muhammad Nor Azmi atas bantuan dan dukungannya selama ini.
8. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna, namun penulis mengharapkan bantuan berupa saran dan kritik yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini.

Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 23 Mei 2023

Penulis

DAFTAR ISI

DAFTAR TABEL

Tabel **Halaman**

DAFTAR GAMBAR

Gambar **Halaman**

DAFTAR LAMPIRAN

Lampiran