

**INTEGRASI METODE K-NN *IMPUTATION* PADA ALGORITMA
DECISION TREE UNTUK KLASIFIKASI PENYAKIT DIABETES**



SKRIPSI

**diajukan untuk memenuhi salah satu syarat
memperoleh gelar sarjana komputer**

oleh :

FATIHUL SUBEQI

2015.69.04.0013

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS YUDHARTA PASURUAN**

2019

PERNYATAAN PENULIS

JUDUL : INTEGRASI METODE K-NN IMPUTATION PADA
ALGORITMA DECISION TREE UNTUK KLASIFIKASI
PENYAKIT DIABETES

NAMA : FATIHUL SUBEQI

NIM : 2015.69.04.0013

"Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut".

Pasuruan, 10 Agustus 2019



Fatihul Subeqi

Penulis

PERSETUJUAN SKRIPSI

JUDUL : INTEGRASI METODE K-NN IMPUTATION PADA
ALGORITMA DECISION TREE UNTUK KLASIFIKASI
PENYAKIT DIABETES
NAMA : FATIHUL SUBEQI
NIM : 2015.69.04.0013

Skripsi ini telah diperiksa dan disetujui

Pasuruan, 10 Agustus 2019

Kepada,

M. Imron Rosadi, S.Kom., M.Kom
NIK. Y 069.02.13.121

Pembimbing,

Moch Latfi, S.Kom., M.Kom
NIK. Y 069.17.09.008

PENGESAHAN SKRIPSI

JUDUL : INTEGRASI METODE K-NN IMPUTATION PADA
ALGORITMA DECISION TREE UNTUK KLASIFIKASI
PENYAKIT DIABETES
NAMA : FATHUL SUBEQI
NIM : 2015.69.04.0013

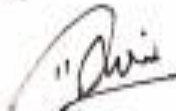
Skripsi ini telah diujikan dan dipertahankan didepan Dewan Penguji pada
Sidang Skripsi tanggal 20 Juli 2019. Menurut pandangan kami, Skripsi ini
memadai dari segi kualitas untuk tujuan penganugerahan gelar
Sarjana Komputer (S.Kom)

Pasuruan, 10 Agustus 2019

Ketua Penguji,


M. Fajishol Amrulloh, S.Kom M.Kom
NIK. Y. 069.17.09.007

Anggota,


Arif Tri Arsanto, S.Kom., M.Kom
NIK. Y. 069.02.01.004

Dekan Fakultas Teknik


Mubach Muir, S.T., M.T
NIK. Y. 069.02.01.015

Pembimbing


Moch Lutfi, S.Kom, M.Kom
NIK. Y. 069.17.09.008

PERSEMBAHAN
BISMILLAHIRROHMANIRROHIM

Atas Berkah Rohmat Allah Yang Maha Kuasa

Janganlah kamu bersikap lemah dan janganlah kamu bersedih hati, padahal kamulah orang yang paling tinggi (derajadnya) jika kamu orang-orang yang beriman (Q.s Ali Imran.39)

Alhamdulillahirobbil Alamin...

Rasa syukur berlimpah hanya kepada Allah...

Manjadda wajadda...

Kata sakti yang telah membuatku bangkit
Meskipun jalan yang ditempuh terjal dan sulit
Tak menyurutkan semangatku walau sedikit

Aku percaya janji Allah pasti

Walau sulit tetap ku jalani

Karena tidak ada yang berharga didunia ini

Selain senyum bangga dibibir orang tuaku

Saat ku persembahkan karya tulis ini...

Karya tulis ini saya persembahkan untuk:

Abah dan Ummi' saya tercinta (H. Toyib & Hj. Bahiyah) terima kasih atas tetesan keringatmu, jerih payahmu, do'amumu yang selalu menyertai langkahku, dan dukungan kalian adalah kekuatan terdahsyat ananda dalam menyelesaikan karya tulis ini.

Dan tak lupa pula kepada adikku tersayang (Alim Maulidin) dan seluruh keluarga yang telah mendo'akan juga semua Guru/Dosen yang telah mendidik saya, Teman-teman yang saya banggakan dan tak lupa untuk para sahabat yang juga telah mensupport saya.

Terima-kasih untuk semua yang telah diberikan kepada saya.

By: Bahie

ABSTRACT

Missing value often occurs in the classification method because information about objects is not provided, difficult to find or does not exist. This leads to decreased data accuracy and quality when data is processed. The classification method used is the K-NN method, because this method includes a classification method that has a strong consistency. This method seeks the case by calculating the closeness between the new case and the old case based on the closest K or Terladder value and can be done to address the missing value, proven by increasing classification and loss of data that has not yet been Classified. The Decision Tree itself for the classification of the dataset but the result is not so accurate then added K-NN method then can be known the value $K = 1$ has a high level of accuracy compared with the value K is another 71.62%, so in this research The K-NN method is 1. Keywords – Missing value, K-nearest neighbor, Decision Tree

Keywords : Missing value, K-nearest neighbor, Decision Tree

ABSTRAK

Missing value sering terjadi dalam metode klasifikasi dikarenakan informasi tentang obyek tidak diberikan, sulit dicari atau memang informasi tersebut tidak ada. Hal ini menyebabkan menurunnya keakuratan dan kualitas data pada saat data diolah. Metode klasifikasi yang digunakan yaitu metode K-NN, karena metode ini termasuk metode klasifikasi yang memiliki konsistensi yang kuat. Metode ini mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama berdasarkan nilai K atau tertangga terdekat dan dapat dilakukan untuk mengatasi *missing value*, terbukti dengan meningkatnya hasil klasifikasi dan hilangnya data yang belum terklasifikasi. *Decision Tree* sendiri untuk klasifikasi dataset namun hasilnya tidak begitu akurat maka ditambahkan metode K-NN kemudian Dapat diketahui nilai $k = 1$ memiliki tingkat akurasi yang tinggi dibandingkan dengan nilai k yang lain yaitu 71.62%, jadi dalam penelitian ini nilai k yang dipakai pada metode K-NN adalah 1.

Kata kunci – *Missing value, K-nearest neighbor, Decision Tree*

KATA PENGANTAR

Dengan mengucapkan syukur alhamdulillah, penulis memanjatkan puji syukur kehadiran Allah SWT, atas rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi ini dengan judul “INTEGRASI METODE KNN IMPUTATION PADA ALGORITMA DECISION TREE UNTUK KLASIFIKASI PENYAKIT DIABETES” sebagai salah satu syarat yang harus dipenuhi untuk memperoleh gelar sarjana Komputer di Universitas Yudharta Pasuruan.

Sholawat dan salam terlimpah selalu kepada junjungan kita Nabi Muhammad SAW sebagai tauladan (Uswatun Hasanah) yang telah menebarkan risalah kebenaran dimuka bumi ini.

Dalam penulisan skripsi ini, Penulis menyadari bahwa banyak hambatan dan kesulitan yang penulis hadapi, alhamdulillah berkat rahmat dan pertolongan Allah SWT serta bantuan dari berbagai pihak akhirnya penulis dapat menyelesaikan skripsi ini. Untuk itu penulis mengucapkan banyak terimakasih kepada semua pihak yang telah membantu serta memberikan dukungan moril maupun materil baik secara langsung ataupun tidak langsung dalam penyusunan skripsi ini. Dalam hal ini penulis mengucapkan terimakasih kepada:

1. Romo KH. M. Sholeh Bahruddin selaku pengasuh Yayasan Darut Taqwa, dimana Universitas Yudharta Pasuruan bernaung dan atas nasehat-nasehat beliau dalam hidup berdampingan dengan masyarakat tanpa membedakan semua kalangan.
2. Teristimewa kepada keluarga tercinta Bapak dan Ibu, yang senantiasa mendoakan dan selalu tak henti-hentinya memberi dorongan semangat dengan penuh kesabaran, penuh kasih dan ketulusan serta perhatiannya telah memberikan support baik materi, moril serta doa yang tidak pernah putus sehingga penulis dapat menyelesaikan skripsi ini.
3. Bapak Dr. Saifulloh, M.HI, Selaku Rektor Universitas Yudharta Pasuruan.
4. Bapak Misbach Munir, ST, MT, Selaku Dekan Fakultas Teknik.
5. Bapak M. Imron Rosadi, S.Kom., M. Kom selaku prodi Teknik Informatika Universitas Yudharta Pasuruan.

6. Bapak Moch. Lutfhi, S.Kom., M.Kom, dosen pembimbing yang telah sudi membantu dalam penyelesaian skripsi ini.
7. Seluruh Bapak dan Ibu Dosen Program Studi Teknik Informatika yang telah memberikan ilmu dan bimbingannya selama masa kuliah.
8. Rekan-rekan Himpunan Mahasiswa Teknik Informatika Universitas Yudharta Pasuruan.
9. Dan semua pihak yang telah membantu penulis dalam penulisan skripsi ini, yang tidak bisa penulis sebutkan satu per satu, terimakasih banyak atas semangat dan dukungannya.

Pasuruan, 27 Agustus 2019

Penulis

DAFTAR ISI

	Halaman
COVER	i
PERNYATAAN PENULIS	ii
PERSETUJUAN SKRIPSI	iii
PENGESAHAN SKRIPSI	iv
PERSEMBAHAN	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terkait	6
2.2 Rangkuman Penelitian Terkait	9
2.3 Landasan Teori	10
2.3.1 Metode <i>K-Nearest Neighbor</i>	10
2.3.2 <i>Decission tree</i>	11
2.3.3 Evaluasi Hasil	13
2.3.4 <i>Confution Matrix</i>	13
2.4 Kerangka Pemikiran	15
BAB III METODE PENELITIAN	16
3.1 Analisa kebutuhan	16
3.2 Metode Penelitian	16
3.3 Pengumpulan data	17
3.4 Pengolahan awal data	18

3.5	Model/metode yang diusulkan	19
3.6	Eksperimen dan pengujian model	19
3.7	Evaluasi dan validasi hasil	19
3.8	Metode Pemilihan Data	19
3.9	Metode Pengumpulan Data	20
3.10	Diabetes	20
3.10.1	Gejala Diabetes	21
3.10.2	Penyebab Diabetes	21
3.11	Metode Analisis Data	22
3.12	Transformasi Data	22
3.13	Pengukuran Kinerja	22
3.14	Perancangan Penelitian	23
3.15	Model yang Diusulkan	23
BAB IV HASIL DAN PEMBAHASAN		25
4.1	Hasil	25
4.2	Perhitungan Manual	25
4.3	Perhitungan Nilai Parameter	25
4.4	Perhitungan <i>K-nearest Neighbor</i>	26
4.5	Hasil Perhitungan Jarak	27
4.6	Hasil dan Evaluasi K-NN dengan <i>Decision tree</i>	28
4.7	Hasil Perhitungan <i>Confusion Matrix</i>	28
4.8	Perbandingan Metode D3 dengan Knn + D3	30
BAB V KESIMPULAN DAN SARAN		31
5.1	Kesimpulan	31
5.2	Saran	31
DAFTAR PUSTAKA		

DAFTAR TABEL

	Halaman
Tabel 2.1 Daftar Penelitian Terkait	9
Tabel 2.2 Confution Matrix	14
Tabel 3.1 Dataset	20
Tabel 4.1 Tampilan Data Training	26
Tabel 4.2 Tampilan Data Testing	26
Tabel 4.3 Tabel Perhitungan Jarak	27
Tabel 4.4 Hasil Confusion Matrix	28
Tabel 4.5 Perbandingan Akurasi	30

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Kerangka Pemikiran	15
Gambar 3.1 Penjelasan tahap Penelitian	17
Gambar 3.2 Model yang diusulkan	23

