

**IMPLEMENTASI METODE MEAN UNTUK MENANGANI
MISSING VALUE PADA ALGORITMA C4.5**



SKRIPSI

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

Oleh :

IMAM GOZALI

2016.69.04.0034

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS YUDHARTA PASURUAN

2020

PERNYATAAN PENULIS

JUDUL : IMPLEMENTASI METODE MEAN UNTUK
MENANGANI MISSNG VALUE PADA
ALGORITMA C4. 5

NAMA : IMAM GOZALI

NIM : 2016.69.04.0034

“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, 3 September 2020



Imam Gozali
Penulis

PERSETUJUAN SKRIPSI

JUDUL : IMPLEMENTASI METODE MEAN UNTUK
MENANGANI MISSING VALUE PADA
ALGORITMA C4.5

NAMA : IMAM GOZALI

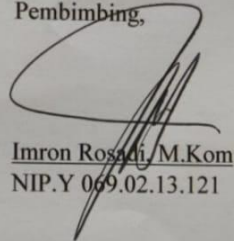
NIM : 2016.69.04.0034

Skripsi ini telah diperiksa dan disetujui
Pasuruan, 3 September 2020

Kaprodi,



Pembimbing,



Imron Rosadi, M.Kom
NIP.Y.069.02.13.121

PENGESAHAN SKRIPSI

JUDUL : IMPLEMENTASI METODE MEAN UNTUK
MENANGANI MISSING VALUE PADA
ALGORITMA C4.5

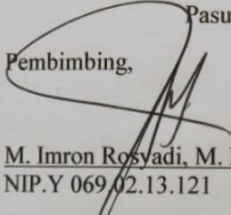
NAMA : IMAM GOZALI

NIM : 2016.69.04.0034

Skripsi ini telah diujikan dan dipertahankan di depan Dewan
Penguji pada Sidang Skripsi tanggal 3 September 2020. Menurut
pandangan kami, Skripsi ini memadai dari segi kualitas untuk
tujuan penganugerahan gelar Sarjana Komputer (S. Kom)

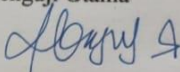
Pasuruan, 3 September 2020

Pembimbing,


M. Imron Rosyadi, M. Kom


NIP.Y 069.02.13.121

Penguji Utama


Cahya Bagus Sanjaya, M.Kom

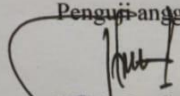
NIP.Y 069.11.13.127

Kaprad,


M. Imron Rosyadi, M.Kom


NIP.Y 069.02.13.121

Penguji-anggota


Moch. Lutfi, M.Kom

NIP.Y 069.16.03.004

Dekan Fakultas Teknik,


Misbach Munir, MT

NIP.Y 069.02.01.015

*“Skripsi ini kutujukan kepada
orang-orang yang berkontribusi
dalam penyelesaian Skripsi ini”*

ABSTRACT

Diabetes Retinopathy is one of the most metabolically in-eye diseases in the world. As time increases and more people have diabetes, there needs to be an early and accurate diagnosis to improve the treatment of diabetes.

In a study many people who have tried several methods to find better result or accuracy one of them is the Decision Tree Algorithm C4.5. In the algorithm C4.5 method will later be added mean method to overcome missing value in Pima Indians Diabetes Database (PIDD).

In the predicted results in the amount of data as much as 768 and 9 attributes one of them attribute target, model decision tree algorithm C4.5 produces 73.5% accuracy while the C4.5 algorithm model plus mean method delivers 76% accuracy.

Keywords : Diabetes, Algorithm C4.5, Mean, Accuracy

ABSTRAK

Diabetes Retinopati adalah salah satu penyakit pada mata yang metaboliknya meningkat secara drastis di dunia. Seiring bertambahnya waktu dan semakin banyaknya orang yang mengalami penyakit diabetes, maka perlu adanya diagnosis lebih dini dan akurat untuk membantu penanganan penyakit diabetes.

Dalam sebuah penelitian banyak orang yang telah mencoba beberapa metode untuk menemukan hasil atau akurasi yang lebih baik salah satunya adalah Decision Tree Algoritma C4.5. Dalam metode algoritma c4.5 nantinya akan ditambahkan metode Mean untuk mengatasi missing value pada data Pima Indians Diabetes Database (PIDD).

Pada hasil prediksi dalam jumlah data sebanyak 768 dan 9 atribut yang salah satunya attribute target, model decision tree algoritma c4.5 menghasilkan akurasi sebesar 73,5% sedangkan model algoritma c4.5 ditambah metode mean menghasilkan akurasi sebesar 76%.

Kata Kunci : Diabetes, Algoritma C4.5, Mean, Akurasi.

KATA PENGANTAR

Alhamdulillah segala puji syukur dipanjatkan atas kehadiran Allah SWT, yang telah memberikan Rahmat, Karunia dan HidayahNya sehingga dapat menyelesaikan penyusunan proposal skripsi ini dengan baik.

Dalam menyusun laporan skripsi ini, penyusun memperoleh bimbingan, pengarahan dan masukan dari berbagai pihak. Penyusun menyadari bahwa proposal skripsi ini tidak akan terselesaikan dengan baik tanpa adanya bantuan mereka. Oleh karena itu, dalam kesempatan ini penyusun mengucapkan terima kasih dengan sebesar-besarnya kepada:

1. Kedua Orang Tua tercinta yang senantiasa mendo'akan serta curahan kasih sayang yang tak terhingga sepanjang masa.
2. Romo Kyai Sholeh Bahruddin, selaku pembimbing dan pembina yang menaungi Universitas Yudharta Pasuruan.
3. Bapak Dr. H. Kholid Murtadlo, SE., ME., selaku Rektor Universitas Yudharta Pasuruan.
4. Bapak Misbach Munir, ST., MT., selaku Dekan Fakultas Teknik Universitas Pasuruan.
5. Bapak Muhammad Imron Rosadi, S.Kom., M. Kom., selaku Ketua Program Studi Teknik Informatika Universitas Yudharta Pasuruan dan sebagai dosen pembimbing yang selalu mendampingi dan memberikan masukan.
6. Teman-teman dan semua pihak yang telah membantu kelancaran pelaksanaan skripsi yang tidak dapat saya sebutkan satu persatu.

Semoga Allah SWT membalas atas seluruh bantuan yang telah dilakukan kepada kami. Aamiin. Kami menyadari bahwa proposal skripsi ini masih jauh dari kesempurnaan. Untuk itu kami harapkan adanya kritik dan saran dari semua pihak demi kesempurnaan dari proposal skripsi ini.

DAFTAR ISI

HALAMAN SAMPUL	i
PERNYATAAN PENULIS	ii
PERSETUJUAN SKRIPSI	iii
PENGESAHAN SKRIPSI	iv
KATA PENGANTAR	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian Terkait	5
2.2 Kajian Teori	7
2.3 Kerangka Pemikiran	13
BAB III METODE PENELITIAN	15
3.1 Metodologi Penelitian	15
3.2 Pengumpulan Data	15
3.3 Pengolahan Data	17
3.4 Metode Usulan	18

BAB IV HASIL DAN PEMBAHASAN	21
4.1 Hasil dan Pembahasan	21
4.1.1 Pengisian Missing Value	21
4.1.2 Preprocessing	22
4.1.3 Penerapan Algoritma C4. 5.....	22
4.1.4 Confusion Matrix.....	28
BAB V KESIMPULAN DAN SARAN	33
5.1 Kesimpulan	33
5.2 Saran	33
LAMPIRAN	36

DAFTAR TABEL

Tabel 3. 1 Dataset	16
Tabel 3. 2 Missing Value	18
Tabel 4. 1 Data Missing Value	21
Tabel 4. 2 Data sesudah dihitung missing value	21
Tabel 4. 3 Nilai Data Minimal dan Maximum	22
Tabel 4. 4 Nilai Data Normalisasi	22
Tabel 4. 5 Nilai Data Entropy	24
Tabel 4. 6 Nilai Data Gain Ratio	25
Tabel 4. 7 Node Tertinggi	26
Tabel 4. 8 Nilai Data Entropy Node 1.1	26
Tabel 4. 9 Nilai Data Gain Ratio Node 1.1	27

DAFTAR GAMBAR

Gambar 2. 1 Kerangka Pemikiran	13
Gambar 3. 1 Metode Usulan.....	18
Gambar 4. 1 Confusion Matrix Pyton C4. 5.....	29
Gambar 4. 2 Confusion Matrix Pyton Mean+C4. 5.....	30
Gambar 4. 3 Grafik Akurasi.....	31

DAFTAR LAMPIRAN

1. Daftar Riwayat Hidup
2. Kartu Seminar
3. Kartu Bimbingan