

**PERBANDINGAN AKURASI ANTARA METODE *K-NEAREST
NEIGHBOR* (KNN) DAN *ARTIFICIAL NEURAL NETWORK*
(ANN) UNTUK KLASIFIKASI INDEKS PEMBANGUNAN
MANUSIA KABUPATEN/KOTA DI PULAU JAWA**

SKRIPSI

**Untuk Memenuhi Persyaratan
Memperoleh Gelar Sarjana**

Oleh:

**Garwita Widyadhana Putri
NIM 19.51.0003**



PROGRAM STUDI S-1 SISTEM INFORMASI

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
PPKIA PRADNYA PARAMITA
MALANG
2023**

LEMBAR PERSETUJUAN

PERBANDINGAN AKURASI ANTARA METODE *K-NEAREST NEIGHBOR* (KNN) DAN *ARTIFICIAL NEURAL NETWORK* (ANN) UNTUK KLASIFIKASI INDEKS PEMBANGUNAN MANUSIA KABUPATEN/KOTA DI PULAU JAWA

Oleh:
Garwita Widyadhana Putri
NIM 19.51.0003

Telah disetujui oleh pembimbing untuk diujikan
Pada tanggal 9 Mei 2023

Menyetujui,
Komisi Pembimbing

Pembimbing I

Dr. Mochamad Husni, S.Pd.I., MM., MMSI
NIK 01.09.10.036

Pembimbing II

Dr. Rahayu Widayanti, S.E., MM., MMSI
NIK 00.12.01.006

Mengetahui,

Ketua Program Studi S-1 Sistem Informasi



Dr. Dwi Safiron Utadina, S.Kom., MMSI
NIK 00.12.01.016

LEMBAR PENGESAHAN

PERBANDINGAN AKURASI ANTARA METODE *K-NEAREST NEIGHBOR* (KNN) DAN *ARTIFICIAL NEURAL NETWORK* (ANN) UNTUK KLASIFIKASI INDEKS PEMBANGUNAN MANUSIA KABUPATEN/KOTA DI PULAU JAWA

Skripsi oleh Garwita Widyadhana Putri
Telah dipertahankan di depan dewan penguji
Pada tanggal 19 Mei 2023

Dewan Penguji,
Ketua

Dr. Dwi Safiroh Utsalina, S.Kom., MMSI
NIK 00.12.01.016

Anggota I

Dr. Weda Adistianaya Dewa, S.Kom., MMSI
NIK 14.02.25.003

Anggota II

Dr. Rahayu Widayanti, S.E., MM., MMSI
NIK 00.12.01.006

Mengetahui,
Ketua Program Studi
S-1 Sistem Informasi



Dr. Dwi Safiroh Utsalina, S.Kom., MMSI
NIK 00.12.01.016

Mengesahkan,
Ketua STMIK PPKIA
Pradnya Paramita



Dr. Tb. Mohammad Akhriza, S.Si., MMSI., Ph.D
NIK 00.12.01.002

**PERNYATAAN
ORISINALITAS SKRIPSI**

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam Naskah SKRIPSI ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu Perguruan Tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah SKRIPSI/TUGAS AKHIR ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia SKRIPSI ini digugurkan dan gelar akademik yang saya peroleh SARJANA dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No 20 Tahun 2003, pasal 25 ayat 2 dan pasal 70).

Malang, 29 Maret 2023

Mahasiswa,

Materai 10000

Nama : Garwita Widyadhana Putri
NIM : 19.51.0003

ABSTRAK

Garwita Widyadhana Putri. 2023. *Perbandingan Akurasi Antara Metode K-Nearest Neighbor (KNN) dan Artificial Neural Network (ANN) untuk Klasifikasi Indeks Pembangunan Manusia Kabupaten/Kota di Pulau Jawa*. Skripsi, Program Studi Sistem Informasi STMIK PPKIA Pradnya Paramita. Pembimbing: (I) Dr. Mochamad Husni, S.Pd.I.,MM.,MMSI, (II) Dr. Rahayu Widayanti, SE.,MM.,MMSI

Kata Kunci : *Akurasi, K-Nearest Neighbor (KNN), Artificial Neural Network (ANN)*

Pengklasifikasian Indeks Pembangunan Manusia (IPM) dilakukan untuk membantu pemerintah dalam mengukur hasil pembangunan sumber daya manusia di Indonesia. Selain itu, hasil IPM juga digunakan pemerintah untuk mengambil kebijakan tertentu pada suatu wilayah, yaitu kabupaten/kota di Pulau Jawa, yang ditinggali sebesar 56,1% dari jumlah penduduk Indonesia. Oleh karena itu, diperlukan pengklasifikasian IPM dengan menggunakan metode *K-Nearest Neighbor* (KNN) dan *Artificial Neural Network* (ANN), untuk mengklasifikasikan IPM kabupaten/kota di Pulau Jawa dengan membandingkan hasil akurasinya. Kedua metode tersebut telah dibandingkan pada penelitian yang berbeda sebelumnya, dengan hasil akurasi yang berbeda pula. Oleh karena itu, penelitian ini akan membandingkan hasil akurasi dari penerapan metode KNN dan ANN pada objek yang sama yaitu IPM kabupaten/kota di Pulau Jawa. Pada penerapan metode KNN, menghasilkan akurasi 95,83% dengan nilai $K=7$ dan pembagian data training dan testing yaitu 80%-20%, namun berdasarkan hasil dari mean *Cross Validation Score* pada metode KNN, yang menghasilkan model terbaik adalah $K=3$. Sedangkan penerapan metode ANN, menghasilkan akurasi 94,44%, pada pembagian data 70%-30%, tetapi bukan model yang terbaik karena *validasi loss* dan *training loss* model ANN dipembagian tersebut termasuk *overfitting*. Berdasarkan hasil penerapan metode KNN dan ANN pada klasifikasi IPM kabupaten/kota di Pulau Jawa menunjukkan bahwa metode KNN lebih unggul secara akurasi, akan tetapi nilai $K=7$ bukan model yang terbaik berdasarkan dari hasil mean *K-Fold Cross Validation Score*, sehingga pada objek penelitian klasifikasi IPM kabupaten/kota di Pulau Jawa tahun 2021, kedua metode tersebut tidak lebih unggul dari satu sama lain.

KATA PENGANTAR

Puji syukur kami panjatkan kehadiran Tuhan Yang Maha Esa atas segala rahmat dan hidayah-Nya peneliti dapat menyelesaikan skripsi yang berjudul **“PERBANDINGAN AKURASI ANTARA METODE *K-NEAREST NEIGHBOR* (KNN) DAN *ARTIFICIAL NEURAL NETWORK* (ANN) UNTUK KLASIFIKASI INDEKS PEMBANGUNAN MANUSIA KABUPATEN/KOTA DI PULAU JAWA”** dengan baik sebagai syarat untuk memperoleh gelar Sarjana Komputer..

Penyusunan SKRIPSI ini tidaklah dapat berjalan dengan baik tanpa adanya dukungan dan kerja sama dari berbagai pihak. Oleh karena itu, rasa terima kasih penulis sampaikan kepada:

1. Tuhan Yang Maha Esa yang telah memberikan petunjuk dan hidayah dalam pembuatan skripsi dan penyusunan laporan sehingga dapat berjalan dengan baik dari awal hingga akhir.
2. Bapak Dr. Tb. Mohammad Akhriza, S.Si., MMSI., Ph.D selaku Ketua STMIK PPKIA Pradnya Paramita dan Dosen Pendamping Akademik.
3. Bapak Dr. Mochamad Husni, S.Pd.I., MM., MMSI selaku dosen pembimbing I
4. Ibu Dr. Rahayu Widayanti, S.E.,MM., MMSI, selaku dosen pembimbing II.
5. Ibu Dr. Dwi Safiroh Utsalina, S.Kom., MMSI selaku Ketua Program Studi S-1 Sistem Informasi.
6. Kedua orang tua, oma, Alm. opa, serta seluruh keluarga besar yang telah memberi semangat dan support yang besar dalam pengerjaan SKRIPSI ini.
7. Teman-teman dekat saya sejak SMP hingga SMA serta teman-teman STMIK PPKIA Pradnya Paramita Malang yang telah memberikan semangat, saran, dan berbagi pengalaman.
8. Dan seluruh pihak yang telah membantu dan mendukung lancarnya pembuatan SKRIPSI dari awal hingga akhir yang tidak dapat disebutkan satu persatu.

Penulis menyadari atas kekurangan dan ketidaksempurnaan dalam penyusunan SKRIPSI ini, sehingga peneliti mengharapkan saran dan kritik dari berbagai pihak yang bersifat membangun demi penyempurnaan SKRIPSI ini.

Semoga SKRIPSI ini berguna bagi pembaca secara umum dan peneliti secara khusus. Akhir kata, penulis ucapkan banyak terima kasih.

Malang, 29 Maret 2023

Penulis,

\

Garwita Widyadhana Putri

DAFTAR ISI

LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN ORISINALITAS SKRIPSI	ii
ABSTRAK	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	5
1.4 Batasan Masalah	5
1.5 Kontribusi Penelitian	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Data Mining	7
2.1.1 Tahap-Tahap Data Mining	9
2.1.2 Pengelompokkan Data Mining.....	11
2.2 Metode Klasifikasi	14
2.3 <i>Artificial Neural Network</i>	16
2.4 <i>K-Nearest Neighbor</i>	20
2.6 <i>K-Fold Cross Validation</i>	22
2.7 Metrik Evaluasi untuk Klasifikasi	23
BAB III METODOLOGI PENELITIAN	25
3.1 Analisis Permasalahan	25
3.2 Solusi yang Diusulkan	26
3.3 Kerangka Kerja (<i>Framework</i>).....	26

3.3.1	Dataset.....	27
3.3.2	<i>Preprocessing</i>	31
3.3.3	Pembagian Data Training, Data Testing, dan Data Validasi	32
3.3.4	Metode <i>K-Nearest Neighbor</i> (KNN).....	34
3.3.5	Metode <i>Artificial Neural Network</i> (ANN)	34
3.3.6	Model Klasifikasi	35
3.3.7	Evaluasi Model.....	35
3.3.8	Perbandingan Hasil Akurasi.....	36
3.4	Rancangan Eksperimen.....	36
BAB IV PENGUJIAN DAN HASIL		38
4.1	Penerapan Metode.....	38
4.1.1	Dataset.....	38
4.1.2	<i>Preprocessing</i>	39
4.1.3	Pembagian Data <i>Training</i> , Data <i>Testing</i> , dan Data Validasi	42
4.1.4	Metode <i>K-Nearest Neighbor</i> (KNN).....	43
4.1.5	Evaluasi Model <i>K-Nearest Neighbor</i> (KNN).....	53
4.1.6	Metode <i>Artificial Neural Network</i>	60
4.1.7	Evaluasi Model <i>Artificial Neural Network</i> (ANN)	65
4.1.8	Perbandingan Hasil Akurasi.....	68
BAB V KESIMPULAN DAN SARAN.....		70
5.1	Kesimpulan	70
5.2	Saran	71
DAFTAR PUSTAKA		72

DAFTAR TABEL

No	Nama Tabel	Halaman
2.1	<i>Confusion Matrix</i>	23
3.1	Sampel Data Umur Harapan Hidup Saat Lahir.....	27
3.2	Sampel Data Rata-Rata Lama Sekolah	28
3.3	Sampel Data Harapan Lama Sekolah.....	28
3.4	Sampel Data Pengeluaran Per Kapita yang Disesuaikan	29
3.5	Sampel Data Keterangan Label Status IPM.....	30
3.6	Model <i>5-Fold Cross Validation</i>	33
4.1	Sampel Data Setelah Transformasi	40
4.2	Hasil Akurasi Metode KNN.....	52
4.3	Hasil <i>Cross Validation Score</i> Metode KNN	53
4.4	<i>Confusion Matrix K=3</i> (60%-40%)	54
4.5	<i>Confusion Matrix K=5</i> (60%-40%)	54
4.6	<i>Confusion Matrix K=7</i> (60%-40%)	55
4.7	<i>Confusion Matrix K=3</i> (70%-30%)	56
4.8	<i>Confusion Matrix K=5</i> (70%-30%)	57
4.9	<i>Confusion Matrix K=7</i> (70%-30%)	57
4.10	<i>Confusion Matrix K=3</i> (80%-20%)	58
4.11	<i>Confusion Matrix K=5</i> (80%-20%)	59
4.12	<i>Confusion Matrix K=7</i> (80%-20%)	59
4.13	Hasil Akurasi Metode ANN.....	65
4.14	<i>Confusion Matrix</i> Pembagian 60%-40%.....	66
4.15	<i>Confusion Matrix</i> Pembagian 70%-30%.....	67
4.16	<i>Confusion Matrix</i> Pembagian 80%-20%.....	68

DAFTAR GAMBAR

No	Nama Gambar	Halaman
2.1	Tahapan Data Mining.....	8
2.2	<i>Flowchart Algoritma Artificial Neural Network</i>	17
2.3	<i>Multi Layer Feedforward Network</i>	20
2.4	<i>Flowchart Algoritma K-Nearest Neighbor</i>	21
3.1	<i>Framework Penelitian</i>	26
3.2	Tahapan <i>Preprocessing</i> Data	31
4.1	<i>Pseudocode Import Library</i>	38
4.2	<i>Pseudocode Input Dataset</i>	38
4.3	<i>Pseudocode Transformasi Data</i>	40
4.4	<i>Pseudocode Pembagian X, y, dan Normalisasi Data</i>	41
4.5	Hasil Normalisasi <i>MinMax Scaler</i>	42
4.6	<i>Pseudocode Pembagian Data Training dan Data Testing 60%-40%</i>	43
4.7	<i>Pseudocode Pembagian Data Training dan Data Testing 70%-30%</i>	43
4.8	<i>Pseudocode Pembagian Data Training dan Data Testing 80%-20%</i>	43
4.9	<i>Pseudocode Model KNN dan K-Fold Cross Validation 60-40%</i>	44
4.10	<i>Pseudocode Model KNN dan K-Fold Cross Validation 70%-30%</i>	45
4.11	<i>Pseudocode Model KNN dan K-Fold Cross Validation 80%-20%</i>	46
4.12	<i>Pseudocode Model KNN dan K-Fold Cross Validation 60%-40%</i>	47
4.13	<i>Pseudocode Model KNN dan K-Fold Cross Validation 70%-30%</i>	48
4.14	<i>Pseudocode Model KNN dan K-Fold Cross Validation 80%-20%</i>	49
4.15	<i>Pseudocode Model KNN dan K-Fold Cross Validation 60%-40%</i>	50
4.16	<i>Pseudocode Model KNN dan K-Fold Cross Validation 70%-30%</i>	51
4.17	<i>Pseudocode Model KNN dan K-Fold Cross Validation 80%-20%</i>	52
4.18	<i>Pseudocode Menambahkan Layer pada Model ANN</i>	60
4.19	<i>Pseudocode Compile dan Melatih Model ANN</i>	61
4.20	<i>Pseudocode Menentukan Training Loss dan Validasi Loss</i>	61
4.21	Grafik <i>Training Loss</i> dan <i>Validasi Loss 60%-40%</i>	62
4.22	Grafik <i>Training Loss</i> dan <i>Validasi Loss 70%-30%</i>	63
4.23	Grafik <i>Training Loss</i> dan <i>Validasi Loss 80%-20%</i>	64

4.24	<i>Pseudocode</i> Hasil Akurasi Pembagian Data 60%-40%	64
4.25	<i>Pseudocode</i> Hasil Akurasi Pembagian Data 70%-30%	64
4.26	<i>Pseudocode</i> Hasil Akurasi Pembagian Data 80%-20%	65
4.27	Grafik Hasil Akurasi Metode <i>K-Nearest Neighbor</i>	68
4.28	Grafik Hasil Akurasi Metode <i>Artificial Neural Network</i>	69

DAFTAR LAMPIRAN

No	Nama Lampiran	Halaman
7.1	Riwayat Hidup	75
7.2	Berita Acara Bimbingan.....	76