

SKRIPSI

PERBANDINGAN METODE K-NEAREST NEIGHBOR DAN

NAÏVE BAYES UNTUK TEXT MINING PADA KLAIM ASURANSI

UNTUK MENDETEKSI POLA RIWAYAT PENYAKIT KRONIS



Disusun Oleh :

ACHMAD VIQIH NURFAUZI

2018230118

PROGRAM STUDI TEKNOLOGI INFORMASI

FAKULTAS TEKNIK

UNIVERSITAS DARMA PERSADA

JAKARTA

2023



UNIVERSITAS DARMA PERSADA

Jl. Taman Malaka Selatan, Pondok Kelapa, Jakarta Timur, Indonesia 13450

Telp. (021) 8649051, 8649053, 8649057 Fax. (021) 8649052

E-mail : humas@unsada.ac.id Home page : <http://www.unsada.ac.id>

LEMBAR PERBAIKAN SIDANG SKRIPSI

Nama : Achmad Viqih Nurfauzi
NIM : 2018230118
Prodi : Teknologi Informasi
Dosen Pembimbing : Bagus Mahardika, S.Kom., MMSI.
Judul : Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Untuk Text Mining Pada Klaim Asuransi Untuk Mendeteksi Pola Riwayat Penyakit Kronis

No	Keterangan	Dosen

Mengetahui,

Kajur Teknik Informatika

Adam Arif Budiman, ST, M.Kom.

MONOZUKURI • TRILINGUAL • ENERGI TERBARUKAN



Lembaga Layanan
Pendidikan Tinggi



TERAKREDITASI B
SK BAN-PT NO. 603/SK/BAN-PT/ANJED/VI/2015



ASOSIASI PENGUSAHA TINGGI SWASTA INDONESIA
APTISI



LEMBAR BIMBINGAN SKRIPSI TEKNOLOGI INFORMASI DARMA PERSADA

NAMA LENGKAP : Achmad Viqih Nurfauzi
NIM : 2018230118
DOSEN PEMBIMBING : Bagus Tri Mahardika, S.Kom, MMSI
JUDUL : Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Untuk Text Mining Pada Klaim Asuransi Untuk Mendeteksi Pola Riwayat Penyakit Kronis

No	Tanggal	Materi	Paraf Dosen Pembimbing
1.	3 Mei 2023	Pengenalan topik skripsi, tujuan penelitian, dan rumusan masalah.	
2.	9 Mei 2023	Pembahasan terkait revisi dari sidang judul.	
3.	16 Mei 2023	Penjelasan tentang dataset yang akan digunakan, termasuk atribut dan kelas yang terkait dengan klaim asuransi dan riwayat penyakit kronis.	
4.	30 Mei 2023	Implementasi metode K-Nearest Neighbor (KNN) untuk text mining pada	

		dataset klaim asuransi dan riwayat penyakit kronis.	
5.	14 Juni 2023	Implementasi metode Naïve Bayes untuk text mining pada dataset klaim asuransi dan riwayat penyakit kronis.	
6.	21 Juni 2023	Evaluasi hasil eksperimen menggunakan metode K-Nearest Neighbor (KNN) dan Naïve Bayes, termasuk pengukuran akurasi dan kinerja.	
7.	04 Juli 2023	Penjelasan laporan skripsi dari bab 1 – bab 4	
8.	12 Juli 2023	Penyusunan bab kesimpulan, saran, dan revisi pada draft skripsi.	
9.	17 Juli 2023	Penjelasan Aplikasi Skripsi dan Laporan Skripsi	

Jakarta, 20 Juli 2023

Dosen Pembimbing,

Bagus Tri Mahardika, S.Kom, MMSI

LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Achmad Viqih Nurfauzi

NIM : 2018230118

Fakultas : Teknik

Jurusan : Teknologi Informasi

Judul : Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Untuk Text Mining Pada Klaim Asuransi Untuk Mendeteksi Pola Riwayat Penyakit Kronis

Menyatakan bahwa Laporan Skripsi ini saya buat sendiri berdasarkan hasil peninjauan, penelitian lapangan, wawancara serta memadukannya dengan berbagai referensi lain yang terkait dan relevan di dalam penyelesaian laporan skripsi ini.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Jakarta, 20 Juli 2023

Achmad Viqih Nurfauzi

LEMBAR PENGESAHAN

PERBANDINGAN METODE K-NEAREST NEIGHBOR DAN NAÏVE BAYES
UNTUK TEXT MINING PADA KLAIM ASURANSI UNTUK MENDETEKSI POLA
RIWAYAT PENYAKIT KRONIS

Disusun oleh :

Nama : Achmad Viqih Nurfauzi

NIM : 2018230118

Mochammad Auditya Brilliant
Pembimbing Lapangan

Bagus Tri Mahardika, S.Kom. MMSI
Pembimbing Laporan

Adam Arif Budiman, S.T., M. Kom.
Ketua Jurusan Teknologi Informasi

LEMBAR PENGUJI SKRIPSI

Laporan Skripsi yang berjudul : “PERBANDINGAN METODE K-NEAREST
NEIGHBOR DAN NAÏVE BAYES UNTUK TEXT MINING PADA KLAIM
ASURANSI UNTUK MENDETEKSI POLA RIWAYAT PENYAKIT KRONIS”

ini telah diajukan pada tanggal

15 Agustus 2023

Penguji 1

Timor Setyaningsih, ST, MTI

Penguji 2

Suzuki Syofian, M.Kom.

Penguji 3

Adam Arif Budiman, ST, M.Kom

ABSTRAK

Asuransi adalah sebuah mekanisme perlindungan terhadap pihak tertanggung apabila mengalami resiko di masa yang akan datang dimana pihak tertanggung akan membayar premi guna mendapatkan ganti rugi dari pihak penanggung. Sejatinnya, asuransi berfungsi sebagai perlindungan jika hal-hal yang tidak diinginkan terjadi. Dalam bisnis asuransi biasanya terdapat 3 pihak yang terkait, perusahaan asuransi, perusahaan broker asuransi dan nasabah asuransi. Salah satu Hal yang menjadi masalah untuk perusahaan asuransi selaku penanggung adalah permasalahan pada tindak kecurangan klaim atau biasa disebut fraud. Salah satu cara yang biasa dilakukan nasabah dalam kasus ini adalah memalsukan jawaban dari riwayat penyakit kronis. Pertanyaan riwayat penyakit kronis ditanyakan kepada nasabah ketika ingin mendaftar asuransi. Banyak kasus dimana nasabah menjawab tidak memiliki riwayat penyakit kronis namun ketika meninggal, penyebab meninggal nya adalah penyakit kronis yang telah diderita lama. Penyebab meninggal ini didapat dari kronologi meninggal yang diisi oleh perwakilan nasabah ketika proses klaim. Untuk saat ini data klaim yang masuk pada sistem PT PANI ada sekitar 10 ribu laporan klaim. Dengan data sebanyak itu tentu saja sangat sulit dan membutuhkan waktu lama untuk mengecek kronologi klaim dan mengklasifikasikan nya dalam keputusan itu penyakit kronis atau tidak. Salah satu teknik yang dapat digunakan untuk menyaring banyaknya data kronologi adalah dengan membuat text mining. Namun, sebelum membuat aplikasi text mining, hal yang perlu dipertimbangkan adalah penerapan metode yang akan dipakai. Dalam penelitian ini, penulis akan membandingkan 2 metode klasifikasi yaitu k-nearest neighbor (KNN) dan naïve bayes classifier. Hasil keluaran dari sistem ini adalah untuk mengklasifikasikan secara otomatis setiap laporan klaim apakah termasuk penyakit kronis atau tidak. Kelebihan dan kekurangan masing-masing metode dalam berbagai situasi. K-Nearest Neighbor terbukti efektif jika fokusnya adalah mengurangi kesalahan positif palsu, sementara Naïve Bayes lebih cocok untuk identifikasi lebih awal. Berdasarkan evaluasi yang dilakukan, Naïve Bayes Classifier ditentukan sebagai metode yang lebih baik dalam konteks ini.

Kata Kunci : Asuransi, Naïve Bayes Classifier, Klaim, K-Nearest Neighbor, Fraud, Koperasi.

KATA PENGANTAR

Puji dan syukur kehadiran Tuhan Yang Maha Esa, atas berkah dan rahmat-Nya, sehingga penulis dapat menyelesaikan laporan skripsi yang berjudul PERBANDINGAN METODE K-NEAREST NEIGHBOR DAN NAÏVE BAYES UNTUK TEXT MINING PADA KLAIM ASURANSI UNTUK MENDETEKSI POLA RIWAYAT PENYAKIT KRONIS. Maksud dan tujuan dari penulisan laporan skripsi ini adalah untuk memenuhi persyaratan kurikulum Sarjana Strata 1 Jurusan Teknologi Informasi di Universitas Darma Persada.

Penulis menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, selama masa perkuliahan sampai pada penyusunan laporan skripsi ini, sangatlah sulit bagi penulis untuk menyelesaikan laporan ini. Maka pada kesempatan ini penulis menyampaikan ucapan terimakasih kepada :

1. Bapak Adam Arif Budiman, S.T., M.Kom., sebagai ketua jurusan teknologi informasi sekaligus pembimbing akademik saya, Universitas Darma Persada.
2. Bapak Bagus Tri Mahardika, S.Kom., M.Si., sebagai dosen pembimbing sekaligus dosen jurusan teknologi informasi, Universitas Darma Persada.
3. Seluruh dosen jurusan teknologi informasi, Universitas Darma Persada.
4. Penulis juga mengucapkan terimakasih kepada kedua orang tua serta keluarga yang telah mendukung penulis dalam pembuatan laporan skripsi ini.
5. Seluruh rekan-rekan mahasiswa Himpunan Mahasiswa Teknologi Informasi Universitas Darma Persada dan Mahasiswa Himpunan lainnya yang tidak bisa penulis sebutkan satu persatu.

6. Rekan-rekan Seluruh Angkatan 2018 Teknologi Informasi Universitas Darma Persada yang tidak bisa penulis sebutkan satu persatu.
7. Rekan-rekan Data Teknologi Terintegrasi khususnya Mochammad Auditya Brilliant, Atika Amalia, Muhammad Januar Fitrianto, Nadia Salsabil dan teman teman saya yang tidak bisa penulis sebutkan satu persatu.

Dengan keterbatasan pengalaman, ilmu maupun Pustaka yang ditinjau, penulis menyadari bahwa laporan Skripsi ini masih banyak kekurangan dan pengembangan lebih lanjut agar benar-benar bermanfaat. Oleh sebab itu, penulis sangat mengharapkan kritik dan saran agar laporan Skripsi ini lebih sempurna serta sebagai masukan bagi penulis untuk penelitian dan penulisan karya ilmiah dimasa yang akan datang.

Akhir kata, penulis berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga laporan skripsi ini membawa manfaat bagi pengembangan ilmu Teknologi Informasi Universitas Darma Persada.

Jakarta, 20 Juli 2023

Achmad Viqih Nurfauzi

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	ii
DAFTAR ISI.....	iv
DAFTAR GAMBAR	vii
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan dan Manfaat.....	4
1.4.1 Tujuan	4
1.4.2 Manfaat	4
1.5 Metodologi Penelitian	5
1.6 Sistematika Penulisan	5
BAB II.....	7
LANDASAN TEORI.....	7
2.1 Penelitian Sebelumnya Yang Berkaitan	7
2.2 Sekilas Tentang Asuransi dan Asuransi Jiwa.....	9
2.3 Sekilas Tentang Naïve Bayes Classifier.....	10
2.3.1 Naïve Bayes Classifier Sebagai Metode Klasifikasi Text Mining...	10
2.4 Sekilas Tentang K-Nearest Neighbor (KNN)	11
2.4.1 K-Nearest Neighbor Sebagai Metode Klasifikasi Text Mining.....	11
2.5 Pemrograman Aplikasi	13

2.5.1	Website.....	13
2.5.2	HTML5	13
2.5.3	PHP	14
2.5.4	PHP Machine Learning.....	15
2.5.5	Laravel.....	16
2.5.6	CSS.....	16
2.5.7	Javascript.....	17
2.6	Pemodelan Sistem Dengan UML	18
2.6.1	Use Case Diagram.....	18
2.6.2	Activity Diagram.....	20
2.6.3	Sequence Diagram	21
2.6.4	Deployment Diagram	22
BAB III		23
ANALISIS DAN RANCANGAN SISTEM		23
3.1	Analisis Permasalahan.....	23
3.2	Perancangan Sistem.....	23
3.2.1	Pemodelan Sistem Dengan UML.....	24
3.2.2	Perancangan Interface	39
3.2.3	Perancangan Database.....	42
3.2.4	Relasi Database	45
BAB IV		46
HASIL DAN PEMBAHASAN.....		46
4.1	Spesifikasi Sistem.....	46
4.1.1	Hardware	46
4.1.2	Software	46
4.2	Hasil Tampilan Website	47

4.2.1	Halaman Login.....	47
4.2.2	Halaman Dashboard.....	47
4.2.3	Halaman List Data Polis.....	48
4.2.4	Halaman Detail Data Polis.....	48
4.2.5	Halaman Laporan Klaim.....	49
4.2.6	Halaman List Data Klaim.....	49
4.2.7	Halaman Detail Data Klaim.....	50
4.2.8	Halaman Konfigurasi User.....	50
4.2.9	Halaman Dashboard Koperasi.....	51
4.3	Evaluasi Hasil Pengujian Aplikasi.....	51
4.3.1	Evaluasi Aplikasi.....	51
4.3.2	Pengujian Aplikasi.....	52
4.4	Implementasi Algoritma Naïve Bayes Classifier (NBC).....	52
4.4.1	Preprocessing Data.....	53
4.4.2	Pembobotan Kata.....	56
4.4.3	Proses Klasifikasi Data Latih.....	57
4.4.4	Proses Klasifikasi Data Uji.....	60
4.5	Pengukuran dan Evaluasi.....	67
4.5.1	Evaluasi Hasil Klasifikasi K-Nearest Neighbor dan Naïve Bayes Classifier.....	67
BAB V.....		71
PENUTUP.....		71
5.1	Kesimpulan.....	71
5.2	Saran.....	72
DAFTAR PUSTAKA.....		73

DAFTAR GAMBAR

Gambar 2. 1 Gambar cara kerja metode KNN.....	12
Gambar 3. 1 Use Case Super Admin	24
Gambar 3. 2 Use Case Asuransi.....	25
Gambar 3. 3 Use Case Koperasi	26
Gambar 3. 4 Activity Diagram Super Admin Login.....	26
Gambar 3. 5 Activity Diagram Superadmin Lihat Dashboard Asuransi	27
Gambar 3. 6 Activity Diagram Superadmin Lihat Dashboard Koperasi	27
Gambar 3. 7 Activity Diagram Lihat List Data Polis.....	28
Gambar 3. 8 Activity Diagram Kelola data user.....	28
Gambar 3. 9 Activity Diagram Lihat data klaim.....	29
Gambar 3. 10 Activity Diagram Melaporkan Klaim	30
Gambar 3. 11 Activity Diagram mengubah status klaim.....	31
Gambar 3. 12 Activity diagram melihat dashboard asuransi	32
Gambar 3. 13 Activity diagram melihat dashboard koperasi.....	32
Gambar 3. 14 Activity diagram mengubah status klaim.....	33
Gambar 3. 15 Activity diagram melihat list data klaim	34
Gambar 3. 16 Activity diagram melihat list data polis	34
Gambar 3. 17 Activity diagram melihat status klaim	35
Gambar 3. 18 Activity diagram melaporkan klaim.....	36
Gambar 3. 19 Sequence diagram login	36
Gambar 3. 20 Sequence diagram mengelola data user	37

Gambar 3. 21 Sequence diagram melihat list dan detail polis	37
Gambar 3. 22 Sequence diagram mengubah status klaim.....	38
Gambar 3. 23 Sequence diagram melaporkan klaim asuransi	38
Gambar 3. 24 Sequence diagram melihat list data klaim.....	39
Gambar 3. 25 Rancangan halaman dashboard	39
Gambar 3. 26 Rancangan list data polis.....	40
Gambar 3. 27 Rancangan konfigurasi user	40
Gambar 3. 28 Rancangan cek status klaim	41
Gambar 3. 29 Rancangan lapor klaim.....	41
Gambar 3. 30 Gambar relasi database.....	45
Gambar 4. 1 Gambar halaman login	47
Gambar 4. 2 Gambar halaman dashboard.....	47
Gambar 4. 3 Gambar halaman list data polis	48
Gambar 4. 4 Gambar detail data polis.....	48
Gambar 4. 5 Gambar lapor klaim.....	49
Gambar 4. 6 Gambar list data klaim	49
Gambar 4. 7 Gambar detail data klaim	50
Gambar 4. 8 Gambar konfigurasi user	50
Gambar 4. 9 Gambar dashboard koperasi	51

DAFTAR TABEL

Tabel 3. 1 Tabel Database Polis.....	42
Tabel 3. 2 Tabel Data Klaim.....	43
Tabel 3. 3 Tabel User.....	44
Tabel 3. 4 Tabel Kronologi.....	44
Tabel 4. 1 Tabel skenario pengujian.....	52
Tabel 4. 2 Tabel data uji naïve bayes classfier.....	53
Tabel 4. 3 Tabel cleaning.....	53
Tabel 4. 4 Tabel case folding.....	54
Tabel 4. 5 Tabel filtering atau stopword.....	54
Tabel 4. 6 Tabel stemming.....	55
Tabel 4. 7 Tabel Tokenizing.....	55
Tabel 4. 8 Tabel pembobotan kata.....	56
Tabel 4. 9 Tabel klasifikasi data latih.....	58
Tabel 4. 10 Tabel klasifikasi data uji.....	60
Tabel 4. 11 Tabel hasil processing.....	60
Tabel 4. 12 Tabel hasil k-nearest neighbor.....	68
Tabel 4. 13 Tabel hasil naïve bayes classifier.....	68
Tabel 4. 14 Tabel perbandingan confusion matriks.....	70

