

PERBANDINGAN NAIVE BAYES CLASSIFIER DENGAN NEAREST NEIGHBOR UNTUK IDENTIFIKASI PENYAKIT MATA

Wildan Budiawan Zulfikar¹, Nur Lukman²

¹Jurusan Teknik Informatika UIN Sunan Gunung Djati Bandung

²Jurusan Teknik Komputer Politeknik Pajajaran ICB Bandung

¹wildan.b@uinsgd.ac.id, ²nur.lukman@poljan.ac.id

Abstrak—*Naive Bayes Classifier dan Nearest Neighbor merupakan algoritma klasifikasi. Kedua algoritma tersebut dapat dikombinasikan dan dibandingkan untuk mendapatkan akurasi dan kecepatan dalam proses identifikasi penyakit mata. Sebuah penyakit mata dapat diidentifikasi dari gejala-gejala yang timbul pada pasien. Sebuah penyakit mata memiliki gejala yang beragam, sebuah penyakit mata juga dapat memiliki gejala yang menyerupai gejala penyakit mata yang lain. Untuk menentukan atau mengidentifikasi jenis penyakit mata yang diderita pasien diperlukan klasifikasi gejala penyakit mata.*

Kata kunci: Identifikasi, Penyakit Mata, Naive Bayes Classifier, Nearest Neighbor

I. PENDAHULUAN

Menurut Kamus Besar Bahasa Indonesia (KBBI) mengidentifikasi adalah menentukan atau menetapkan identitas (orang, benda, dan sebagainya). Penyakit adalah suatu keadaan abnormal dari tubuh atau pikiran yang menyebabkan ketidaknyamanan, disfungsi atau kesukaran terhadap orang yang dipengaruhinya.

Salah satu usaha untuk mengetahui jenis Penyakit yang diderita pasien dan menyembuhkan suatu Penyakit adalah dengan cara pergi ke dokter dan meminum obat. Tetapi kenyataannya seorang dokter tidak selalu berada di tempat praktek atau di klinik untuk merawat pasien. Sedangkan pasien atau penderita Penyakit tidak mengetahui jenis Penyakit dan obat yang harus dikonsumsi. Pada dasarnya, sebuah Penyakit dapat diketahui atau diidentifikasi jenisnya dilihat dari gejala-gejala yang diderita pasien.

Naive Bayes Classifier adalah metode pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Naive Bayes Classifier (NBC) terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar [1]. Nearest Neighbor merupakan algoritma pembandingan yang digunakan untuk identifikasi penyakit mata.

Naive Bayes Classifier dan Nearest Neighbor akan dikombinasikan dalam rangka memperoleh hasil identifikasi yang akurat. Adapun beberapa aspek yang menjadi bahan atau kriteria dari kedua algoritma tersebut yakni kecepatan dalam proses identifikasi.

II. TINJAUAN PUSTAKA

A. Konsep Klasifikasi

Dalam ajaran Islam, tentu sangat kenal dengan kata klasifikasi atau penggolongan. Misalnya, penggolongan manusia, ibadah, Al-Qur'an dari segi turunnya ayat, ilmu sampai dengan cara penyampaian ilmunya. Sebab dengan adanya klasifikasi atau penggolongan tentu akan mempermudah dalam pelaksanaannya [2].

Sedangkan dalam ilmu pengetahuan dan teknologi, klasifikasi merupakan fungsi mayor dari data mining. Algoritma klasifikasi menggunakan data latihan untuk menghasilkan pengetahuan. Pada klasifikasi, sebuah catatan (*record*) akan diklasifikasi ke dalam salah satu dari sekian klasifikasi yang tersedia pada *variable* tujuan berdasarkan nilai-nilai variabel prediktornya [3].

Terdapat dua teknik pembelajaran (latihan), yaitu *unsupervised* dan *supervised learning*. Pada *supervised learning* terdapat label atau nama kelas pada data latihan (*supervisi*) dan data baru diklasifikasikan berdasarkan data latihan. Sedangkan pada *unsupervised learning* tidak terdapat label atau nama kelas pada data latihan, data latihan dikelompokkan berdasarkan ukuran kemiripan pada suatu kelas.

Berdasarkan keluaran dari fungsi, *supervised learning* dibagi menjadi 2, regresi dan klasifikasi. Regresi terjadi jika *output* dari fungsi merupakan nilai yang berlanjut, sedangkan klasifikasi terjadi jika keluaran dari fungsi adalah nilai tertentu dari suatu atribut tujuan (tidak berlanjut). Tujuan dari *supervised learning* adalah untuk memprediksi nilai dari fungsi untuk sebuah data masukan yang sah setelah melihat sejumlah data latihan.

B. Klasifikasi Penyakit Mata

Penyakit adalah keadaan tidak normal pada badan yang menyebabkan ketidakselarasan, disfungsi, atau tekanan/stres kepada orang yang menderitanya. Terdapat berbagai jenis Penyakit Mata yang mengancam manusia. Penyakit Mata dapat disebabkan oleh kuman, bakteri, virus, racun, kegagalan organ tubuh, dan juga dapat bersumber dari turunan (Penyakit Mata turunan).

Mata merupakan organ yang penting dalam kehidupan manusia. Oleh itu, organ ini perlu dijaga dengan sebaik-baiknya.

Terdapat berbagai jenis Penyakit Mata, dari anak-anak sampai dewasa dan tua. Penyakit Mata ada yang harus segera diobati supaya mata tidak bertambah parah dan bisa menyebabkan kebutaan. Ada juga yang hanya

sekedar kurang jelas penglihatannya yang hanya memerlukan memakai kaca mata untuk memperbaiki penglihatannya.

Setiap Penyakit Mata memiliki ciri-ciri atau sering kita sebut sebagai gejala Penyakit Mata. Setiap Penyakit Mata memiliki gejala yang berbeda-beda namun beberapa Penyakit Mata memiliki gejala yang beragam dan beberapa Penyakit Mata juga memiliki gejala yang mirip.

Pada proses pemeriksaan, Seorang dokter dapat menentukan Penyakit Mata yang diderita pasiennya melalui diagnosa. Pasien menjelaskan apa saja yang menjadi keluhannya. Dokter bisa mengklasifikasi penjelasan dari pasien tersebut, setelah terbentuk sebuah pola gejala Penyakit Mata, dokter akan membandingkan dengan gejala yang sama yang pernah ada sebelumnya. Setelah itu, dokter dapat menyimpulkan Penyakit Mata yang diderita pasien.

Berdasarkan analogi proses berobat di atas, dapat disimpulkan bahwa setiap Penyakit Mata memiliki pola gejala. Pola gejala ini dapat dibandingkan dengan pola gejala baru jika cocok atau memenuhi nilai-nilai tertentu maka pola baru tersebut adalah gejala dari Penyakit Mata yang sama.

Maka untuk mendeteksi Penyakit Mata hal yang pertama dilakukan adalah dengan mengklasifikasi gejala berdasarkan objeknya. Kebanyakan Penyakit Mata muncul pada objek-objek sebagai berikut: penglihatan, kemampuan mata, stabilitas mata, kepekaan terhadap cahaya, kepekaan terhadap warna, keadaan fisik mata, keadaan non fisik mata, kornea, kelopak mata, kantung air mata, dan komplikasi terhadap daya tahan tubuh.

Setiap objek gejala Penyakit Mata memiliki kemungkinan gejala yang timbul. Misalkan pada objek stabilitas mata yang memiliki kemungkinan seseorang mengalami mudah lelah. Mudah lelah adalah gejala yang timbul pada stabilitas mata. Begitu juga dengan objek yang lain memiliki lebih dari satu kemungkinan gejala. Kesebelas objek gejala Penyakit Mata ini akan membentuk pola baru gejala Penyakit Mata. Setelah pola gejala Penyakit Mata terbentuk, pola ini akan dibandingkan dengan pola gejala Penyakit Mata yang sudah ada menggunakan metode statistik.

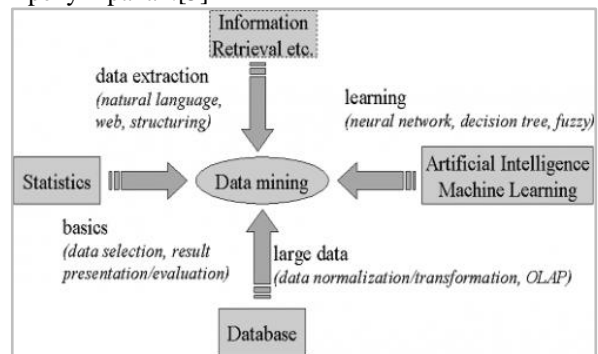
C. Data Mining

“Data mining merupakan analisis dari peminjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Selain itu, data mining juga merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar.[5]”

Larose menyatakan bahwa kemajuan yang luar biasa yang terus berlanjut dalam bidang data mining didorong oleh beberapa faktor, diantaranya adalah:

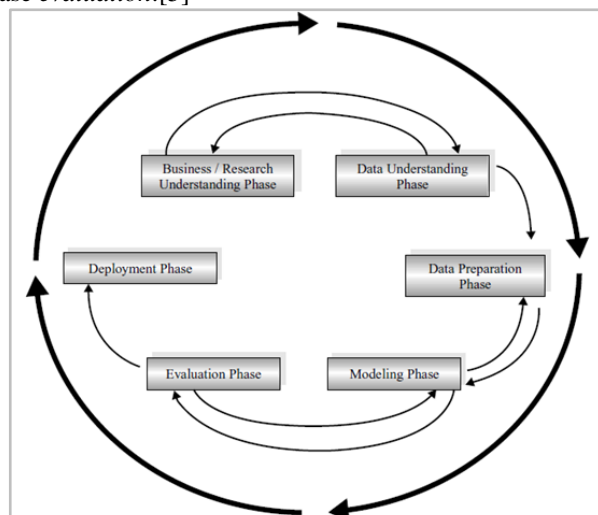
1. Pertumbuhan yang cepat dalam kumpulan data.

2. Penyimpanan data dalam data warehouse, sehingga seluruh perusahaan memiliki akses ke dalam database yang handal.
3. Adanya peningkatan akses data melalui navigasi web dan intranet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk data mining.
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.[5]



Gambar 1 Bidang Ilmu Data Mining

Dalam CRISP-DM, sebuah proyek data mining memiliki siklus hidup yang terbagi dalam enam fase. Keseluruhan fase berurutan yang ada tersebut bersifat adaptif. Fase berikutnya dalam urutan bergantung kepada keluaran dari fase sebelumnya. Hubungan penting antar fase digambarkan dengan panah. Sebagai contoh, jika proses berada pada fase modeling. Berdasarkan pada perilaku dan karakteristik model, proses mungkin harus kembali kepada fase data preparation untuk perbaikan lebih lanjut terhadap data atau berpindah maju kepada fase evaluation.[5]



Gambar 2 Proses Data Mining menurut CRISP-DM

Menurut Larose data mining terbagi menjadi enam fungsi diantaranya adalah fungsi deskripsi (description), fungsi estimasi (estimation), fungsi prediksi (prediction), fungsi klasifikasi (classification), fungsi pengelompokan (classification), dan fungsi asosiasi (association)

[5]. Menurut Berry dan Browne keenam fungsi *data mining* tersebut dipilah menjadi:

- a. Fungsi minor atau fungsi tambahan, yang meliputi ketiga fungsi pertama yaitu deskripsi, estimasi, dan prediksi.
- b. Fungsi mayor atau fungsi utama, yang meliputi ketiga fungsi berikutnya yakni klasifikasi, pengelompokan, dan asosiasi [6].

D. Naïve Bayes Classifier

Data mining memiliki hakikat (*notion*) sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki [3].

Data mining mengubah data tidak terstruktur menjadi data terstruktur melalui tahap-tahap berikut:

1. *Text Preprocess* yaitu pemecahan sekumpulan karakter ke dalam kata-kata
2. *Feature Generation* yaitu mengubah kata-kata ke dalam bentuk dasar sekaligus mengurangi jumlah kata-kata tersebut.
3. *Feature Selection* yaitu seleksi *feature* untuk mengurangi dimensi dari suatu kumpulan teks.
4. *Pattern Discovery* yaitu dapat berupa *unsupervised learning (clustering)* atau *supervised learning (classification)*.
5. *Interpretation/Evaluation* yaitu pengukuran efektivitas untuk mengevaluasi metode yang diterapkan menggunakan parameter *precision*.

Naïve Bayes Classifier atau sering disebut Bayesian Classification adalah metode pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*. Naïve Bayes Classifier (NBC) terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar [1].

Naïve Bayes Classifier didasarkan kepada teorema Bayes yang memiliki bentuk umum sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

Dalam hal ini:

- X = data dengan class yang belum diketahui
- H = hipotesis data X merupakan suatu *class* spesifik
- P(H|X) = probabilitas hipotesis H berdasar kondisi X (*posteriori probability*)
- P(X|H) = probabilitas X berdasar kondisi pada hipotesis H
- P(X) = probabilitas dari X [Kusrini, 2009]

Pada kasus pendeteksian Penyakit Mata metode Naive Bayes Classifier berperan dalam proses perbandingan pola baru gejala penyakit mata yang telah terbentuk dengan gejala umum penyakit mata. Setiap objek Penyakit Mata pada pola baru akan dibandingkan dengan gejala umum Penyakit Mata kemudian dihitung juga berapa besar kemungkinan setiap objek gejala Penyakit Mata menyebabkan setiap Penyakit Mata.

Setiap nilai yang terbentuk diakumulasikan berdasarkan Penyakit Mata. Proses ini menghasilkan nilai bobot untuk setiap jenis Penyakit Mata. Setelah setiap Penyakit Mata memiliki nilai bobot, tahap terakhir adalah memutuskan atau menentukan jenis Penyakit Mata berdasarkan pola baru dengan menetapkan Penyakit Mata yang memiliki nilai bobot terbesar sebagai Penyakit Mata yang diderita pasien.

E. Nearest Neighbor

Nearest Neighbor adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. Misalkan diinginkan untuk mencari solusi terhadap seorang pasien baru dengan menggunakan solusi dari pasien terdahulu. Untuk mencari kasus pasien mana yang akan digunakan, maka dihitung kedekatan kasus pasien baru dengan semua kasus pasien lama. Kasus pasien lama dengan kedekatan terbesar yang akan diambil solusinya untuk digunakan pada kasus pasien baru [1].

$$\text{similarity} = \frac{\sum_{i=1}^n f(T_i, S_i)}{W_i}$$

Keterangan:

- T : Kasus baru
- S : kasus yang ada dalam penyimpanan
- N : jumlah atribut
- i : atribut individu antara 1 – n
- f : fungsi *similarity* ATRIBUT I antara kasus T dan kasus S
- w : bobot yang diberikan pada atribut ke-i

III. FASE-FASE DATA MINING

A. Business Understanding Phase

Penyakit mata diidentifikasi dengan menggunakan Naive Bayes Classifier dan Nearest Neighbor sebagai algoritma perbandingan. Atribut yang terdapat di dalam *training dataset* yakni penglihatan, kemampuan, daya tahan mata, kepekaan terhadap cahaya, kepekaan terhadap warna, kondisi fisik mata, kondisi non fisik mata, kornea mata, kelopak, kantung air mata, dan komplikasi.

Apabila mata pasien tidak menunjukkan adanya gejala-gejala yang tidak normal, maka mata pasien tersebut dinyatakan sehat. Namun apabila pasien menunjukkan gejala yang tidak normal, aplikasi menganalisa gejala-gejala timbul pada pasien tersebut dengan cara membandingkannya dengan gejala umum penyakit mata (*training dataset*).

Proses komparasi dilakukan dengan menggunakan metode Naive Bayes Classifier dan Nearest Neighbor. Proses komparasi antara gejala yang timbul pada mata pasien dan gejala-gejala umum penyakit mata menghasilkan nilai bobot, sehingga setiap Penyakit Mata

yang telah diidentifikasi dan didefinisikan memiliki satu nilai bobot.

Naive Bayes Classifier menentukan jenis Penyakit Mata yang diderita pasien dengan melihat nilai bobot Penyakit Mata. Penyakit Mata yang memiliki nilai bobot terbesar adalah Penyakit Mata yang diderita oleh pasien.

B. Data Understanding and Preparing Phase

Langkah pertama yang dilakukan dalam proses pendeteksian Penyakit Mata menggunakan metode Naive Bayes Classifier adalah menganalisis gejala Penyakit Mata yang timbul pada pasien. Metode Naive Bayes Classifier memerlukan data yang telah berbentuk pola yang berisi gejala yang timbul pada pasien. Karenanya gejala-gejala yang timbul pada mata pasien harus diklasifikasi terlebih dahulu.

C. Modeling Phase

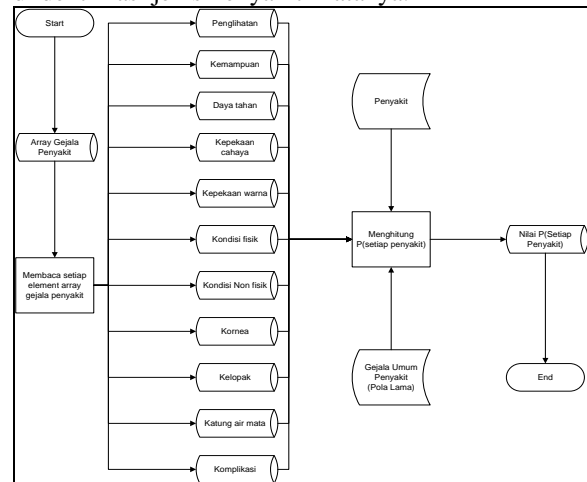
Gejala-gejala Penyakit Mata yang timbul pada pasien akan dipecah atau diklasifikasi ke dalam beberapa objek yakni penglihatan, kemampuan mata, daya tahan mata, kepekaan cahaya, kepekaan warna, kondisi fisik mata, kondisi non fisik mata, kornea, kelopak, kantung air mata, dan komplikasi. Objek-objek tersebut dipilih dengan asumsi bahwa gejala umum setiap Penyakit Mata akan timbul pada satu atau lebih objek-objek tersebut. Suatu objek tempat munculnya gejala suatu Penyakit Mata memiliki indikasi yang berbeda satu sama lain dan satu objek dapat memiliki indikasi lebih dari satu.

Gejala-gejala yang timbul pada pasien akan dicocokkan dengan data indikasi, sehingga diperoleh pola baru berupa kumpulan gejala Penyakit Mata dalam bentuk larik (array) yang siap untuk diproses pada tahap berikutnya. Pada gambar 4.2 dijelaskan bahwa pasien menjelaskan gejala Penyakit Mata yang timbul pada matanya melalui interface yang telah disediakan, lalu gejala tersebut diklasifikasi berdasarkan objek dan indikasinya.

Langkah berikutnya adalah menyiapkan pola gejala Penyakit Mata yang baru untuk dibandingkan dengan gejala umum penyakit mata (Pattern Preparing). Gejala Penyakit Mata yang telah diklasifikasi di-insert ke dalam bentuk array (larik) lalu setiap element array tersebut diperiksa. Bila setiap element array Gejala Penyakit Mata menunjukkan hasil yang normal maka pasien dinyatakan sehat. Jika tidak demikian, array gejala Penyakit Mata yang telah terbentuk akan diproses untuk menentukan jenis Penyakit Mata yang diderita pasien.

Jika kita mendefinisikan sebuah kasus dimana pasien memiliki gejala yang tidak normal pada proses patten preparing, makalangkah berikutnya adalah menghitung nilai bobot setiap Penyakit Mata berdasarkan gejala yang timbul pada pasien (pola baru) dan gejala umum penyakit mata dengan menggunakan metode Naive Bayes Classifier. Pada proses ini setiap element dari array Gejala Penyakit Mata akan dibandingkan dengan gejala umum penyakit mata dengan menghitung berapa banyak indikasi yang muncul pada pasien yang dapat menjadi gejala suatu Penyakit Mata. Gambar 4.4 menjelaskan

proses penghitungan setiap gejala (indikasi) yang timbul pada pasien dibandingkan dengan gejala-gejala yang telah diidentifikasi jenis Penyakit Matanya.



Gambar 3 Flow Chart Naive Bayes Classifier

Nilai-nilai dari seluruh objek akan diakumulasikan berdasarkan jenis Penyakit Mata sehingga diperoleh nilai bobot setiap (tabel 4.4). Langkah terakhir adalah mencari Penyakit Mata yang memiliki nilai bobot terbesar sebagai jenis Penyakit Mata yang diderita pasien. D035 memiliki nilai bobot yang paling besar maka D035 adalah Penyakit Mata yang diderita pasien.

Tabel 1 Hasil Deteksi Menggunakan Naive Bayes Classifier

	D030	D031	D032	D033	D034	D035	D036	D037	D038	D039	DO40
Disease	0,025	0,051	0,025	0,025	0,051	0,025	0,102	0,025	0,051	0,025	0,051
Penglihatan	0	0	0	0	0,5	1	1	0	1	0	0,25
Ability	2	1	1	1	1	1	1	1	1	1	1
Stability	1	1	1	1	1	1	1	1	1	0	1
Pekacahaya	0	1	0	1	1	1	0	0	1	0	1
Pekawarna	1	0	1	1	1	1	1	1	1	1	1
Fisik	0	0	1	1	0	1	0,5	0	0	0	1
Nonfisik	1	1	0	1	1	1	0	1	0	1	1
Kornea	1	1	1	1	0	1	1	0	0	1	1
Kelopak	1	1	1	1	1	1	1	0	1	1	1
Airmata	1	1	1	1	1	1	1	1	0	1	1
Komplikasi	1	1	1	1	1	1	1	0	1	1	1
Total (*)	0	0	0	0	0	0,025	0	0	0	0	0,01275

Proses ini sangat rentan terhadap kesalahan, bila salah-satu objek tidak terakumulasi dalam proses penghitungan nilai bobot akibatnya adalah nilai bobot yang telah dihitung tidak akurat dan dapat menyebabkan kesalahan dalam penentuan jenis penyakit mata atau menyebabkan tidak teridentifikasinya penyakit mata yang diderita pasien.

Dalam penelitian ini diimplementasikan algoritma Nearest Neighbor sebagai algoritma pembanding. Langkah-langkah dalam Nearest Neighbor yakni menentukan bobot setiap atribut, menentukan nilai kedekatan setiap value dalam atribut.

Atribut yang digunakan dalam proses modeling sebanyak sebelas atribut yakni: penglihatan, ability, stability, peka cahaya, peka warna, fisik, non fisik, kornea, kelopak, kantung air mata, komplikasi.

Setelah mengetahui nilai bobot setiap atribut, langkah selanjutnya yakni menentukan nilai kedekatan setiap nilai dalam atribut. Setelah mendapatkan nilai bobot setiap atribut dan nilai kedekatan nilai pada atribut langkah selanjutnya yakni melakukan perhitungan terhadap semua *row* pada *training dataset*.

Setelah dilakukan penghitungan nilai kedekatan pada setiap baris pada *training dataset*, penyakit dengan id D035 dan nama Retinoblastoma memiliki nilai kedekatan yang paling besar.

Tabel 2 Rekap Nilai Kedekatan

No.	Id Penyakit	Penyakit	Nilai Kedekatan
1	D019	Ablasio	0.932
2	D030	Kebutaan	0.909
3	D030	Kebutaan	0.932
4	D031	Buta Warna	0.932
5	D032	Iridosiklitis	0.932
6	D033	Infeksi Mata	0.977
7	D034	Katarak	0.955
8	D034	Katarak	0.932
9	D035	Retinoblastoma	1.000
...

D. Evaluation Phase

Berdasarkan hasil pemodelan menggunakan algoritma Naive Bayes Classifier yang tersaji pada tabel 1 dan hasil pemodelan menggunakan algoritma Nearest Neighbor pada tabel 2, Naive Bayes Classifier dan Nearest Neighbor menghasilkan hasil klasifikasi yang sama. Pada kasus baru, kecenderungan penyakit mata yang diderita pada kasus baru tersebut adalah penyakit dengan id D035 dan nama Retinoblastoma.

Ditinjau dari kecepatan proses, baik Naive Bayes Classifier ataupun Nearest Neighbor memiliki rata-rata selisih waktu yang tidak terlalu jauh yakni Nearest Neighbor lebih cepat 0.027 detik daripada Naive Bayes Classifier.

Tabel 3Perbandingan Kecepatan Algoritma

Test Attemp	Naive Bayes Classifier (s)	Nearest Neighbor (s)
T001	0.153	0.122
T002	0.132	0.130
T003	0.179	0.161
T004	0.190	0.125
T005	0.337	0.150
T006	0.156	0.143
T007	0.122	0.089
T008	0.176	0.160

T009	0.172	0.134
T010	0.088	0.225
Average	0.171	0.144

IV. KESIMPULAN

Algoritma Naive Bayes Classifier dan Nearest Neighbor memiliki keakuratan yang sebanding. Dalam hal kecepatan, Nearest Neighbor memiliki rata-rata catatan lebih cepat 0.027 detik dari pada Naive Bayes Classifier seperti yang disajikan pada tabel 3.

Naive Bayes Classifier dan Nearest Neighbor dapat dikombinasikan dalam hal pengidentifikasian penyakit mata dalam rangka meminimalisir terjadinya kesalahan dalam proses klasifikasi.

Mengkombinasikan algoritma Naive Bayes Classifier dan Nearest Neighbor berdasarkan gejala-gejala yang timbul pada tubuh pasien cukup efektif dalam mengidentifikasi penyakit mata. Model ini dapat dioptimalkan dengan menambahkan teknologi pengolahan citra dengan tujuan meningkatkan akurasi dalam pengidentifikasian penyakit mata.

REFERENSI

- [1] Kusriani., Luthfi, Emha Taufik. 2009. *Algoritma Data Mining*. Yogyakarta: Penerbit Andi
- [2] Musthafa, Aziz. 2009. *Klasifikasi Otomatis Dokumen Berita Kejadian Berbahasa Indonesia*. Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
- [3] Susanto, Sani, Ph.D dan Dedy Suryadi, S.T.,M.S. 2010. *Pengantar Data Mining: Menggali Pengetahuan dari Bongkahan Data*. Yogyakarta: Andi Publisher
- [4] Ilyas, Sidarta. 2009. *Ikhtisar Ilmu Penyakit Mata*. Jakarta: Balai Penerbit FKUI
- [5] Larose, Daniel T. 2005. *Discovering Knowledge In Data: An Introduction to Data Mining*. New Jersey: John Wiley and Sons Inc.
- [6] Berry, M.W. dan M. Browne. 2006. *Lecture Notes in Data Mining*. Singapore: World Scientific Publishing Co. Pte. Ltd.