

BAB II

LANDASAN TEORI

2.1 Mata

Mata merupakan salah satu dari lima panca indra yang manusia miliki, mata merupakan organ penglihatan. Mata merupakan organ penglihatan yang dimiliki manusia. Mata dilindungi oleh area orbit tengkorak yang disusun oleh berbagai tulang seperti tulang *frontal*, *sphenoid*, *maxilla*, *zygomatic*, *greater wing of sphenoid*, *lacrimal*, dan *ethmoid* (Rizzo, 2001).

2.1.1 Bagian-Bagian Mata

Secara garis besar mata memiliki dua bagian utama yaitu bagian dalam bola mata dan bagian luar bola mata.

Bagian luar mata terdiri dari alis mata, kelopak mata, bulu mata. Berikut akan dijelaskan fungsi dari bagian luar mata yaitu sebagai berikut. (Rizzo, 2001).

1. Alis Mata
2. Kelopak Mata
3. Bulu Mata

Bagian dalam mata terdiri dari dinding bola mata, retina, iris, pupil, lensa, kelenjer lakrima, saraf optic, titik buta. Fungsi dari tiap bagian dalam mata akan dijelaskan sebagai berikut:

1. Dinding bola mata
2. Retina
3. Iris
4. Pupil
5. Lensa
6. Kelenjer lakrima (kelenjer air mata)
7. Saraf Optic

2.1.2 Penyakit mata

Penyakit mata adalah segala gangguan atau penyakit yang membuat mata terasa sakit sehingga akan menurunnya daya penglihatan seseorang, kurangnya

focus penglihatan dan dapat juga menyebabkan kelainan pada mata. Penyakit mata sering diakibatkan iritasi atau peradangan akibat infeksi di bagian selaput yang melapisi mata. Penyakit mata tidak bisa dianggap remeh karena mata merupakan indera penting untuk melihat, dengan adanya penyakit mata dapat menyebabkan turunnya produktifitas seorang dalam berkerja (Ilyas S, 2006).

Berikut beberapa jenis penyakit mata antara lain sebagai berikut:

1. Katarak

Katarak adalah sebuah penyakit mata yang menyebabkan lensa mata menjadi keruh dan menyebabkan penglihatan menjadi berkurang yang diakibatkan oleh lendir atau lapisan putih yang menutupi lensa mata. Apabila dibiarkan katarak akan dapat menyebabkan kebutaan. Biasanya katarak terjadi pada orang yang sudah berlanjut usia. Penderita katarak dapat disembuhkan dengan cara mengangkat lensa yang buram tersebut.

2. *Conjunctivitis*

Conjunctivitis atau lebih dikenal dengan mata merah adalah peradangan yang terjadi pada konjungtiva atau selaput bening yang melapisi bagian depan mata. Peradangan biasanya disebabkan oleh infeksi karena bakteri maupun virus dan alergi.

3. *Myopia*

Myopia adalah rabun jauh dimana kondisi penglihatan tidak mampu melihat benda dalam jarak jauh, tapi bisa melihat benda dalam jarak dekat. Gejala awal *Myopia* biasanya seperti penglihatan buram ketika melihat benda dalam jarak jauh, sering menyipitkan mata ketika melihat, sakit kepala dan kesulitan melihat saat malam hari.

4. *Presbiopia*

Presbiopia adalah kondisi mata yang akan dialami manusia saat proses penuaan. Presbiopia merupakan kondisi mata yang kehilangan kemampuan focus secara bertahap untuk melihat suatu objek pada jarak dekat. Gejala presbiopia seperti kebiasaan menyipitkan mata, butuh

lampu lebih terang saat membaca, kesulitan melihat huruf kecil dan penglihatan kabur ketika membaca pada jarak normal.

5. Pterygium

Pterygium adalah penyakit mata yang disebabkan karena adanya pertumbuhan jaringan berbentuk segitiga dilapisan membran tipis bening dibagian putih mata. Pterygium ringan dapat diobati dengan obat tetes mata tetapi apabila gejala yang dirasakan berat maka pengobatan dilakukan dengan operasi pengangkatan.

6. Astigmatisme *Myopia*

Astigmatisme *Myopia* penyakit mata yang biasanya terjadi karena kelainan sejak lahir, cedera pada mata atau akibat operasi mata. Astigmatisme *Myopia* merupakan gangguan penglihatan akibat kelainan pada kelengkungan kornea atau lensa. Pada Astigmatisme *Myopia* ini menyebabkan pandangan kabur baik melihat dalam jarak dekat maupun jarak jauh.

7. Trauma

Trauma mata adalah rusaknya jaringan pada bola mata, kelopak mata, saraf mata dan rongga orbital akibat dari adanya benda asing yang mengenai mata dengan keras secara cepat maupun lambat. Penyebab trauma mata bisa bermacam-macam seperti tergores ranting pohon, benda-benda tajam ataupun tumpul yang mengenai mata dan paparan bahan kimia.

8. *Dry eye*

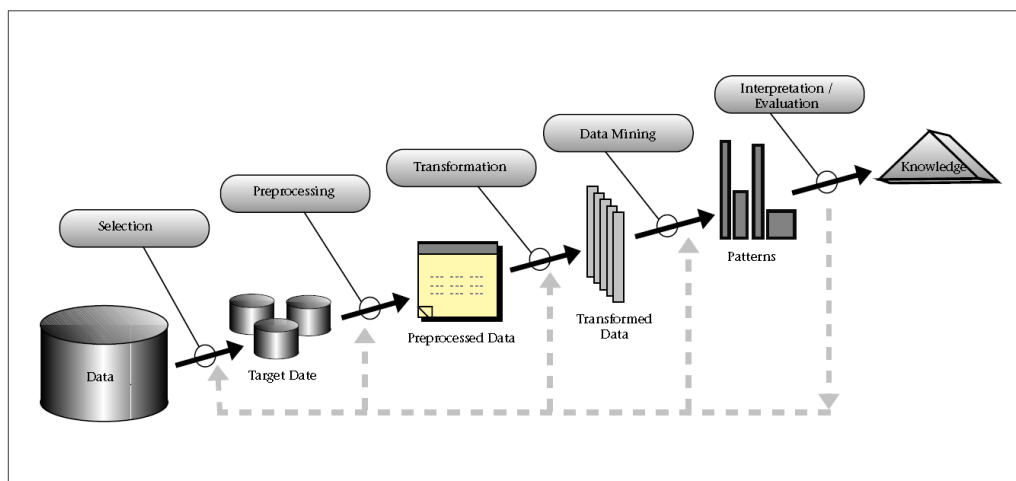
Dry eye merupakan kondisi dimana mata menjadi kering yang disebabkan jumlah air mata yang tidak stabil dan kurang berkualitas air mata sehingga kemampuan dalam melindungi dan melembabkan permukaan mata menjadi berkurang. Gejala *dry eye* seperti mata terasa perih, gatal dan kering pada mata, mata mudah merah, mata sering mengeluarkan kotoran dan mata terasa berat dan pegal. Dalam *dry eye* bisa terjadi kondisi *reflex tearing* yaitu kondisi mata akan terus berair tetapi tidak mampu mengatasi mata yang kering.

2.2 Data Mining

Data mining adalah analisa otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya. Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda sehingga dapat dipahami dan bermanfaat bagi pemilik data (Larose, 2005).

2.3 Knowledge Discovery in Database (KDD)

Menurut mujiasih (2011) Data mining atau sering disebut *knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan data, pemakaian data histori yang bertujuan untuk menemukan keteraturan, pola atau hubungan dalam data yang berukuran besar.



Gambar 2.1 Tahapan- Tahapan KDD (Sumber : Fayyad dkk, 1996)

Berikut beberapa tahapan dari *Knowledge Discovery in Database* (KDD) yaitu sebagai berikut:

1. Data Selection

Dalam tahapan *data selection* atau seleksi ini, memiliki fungsi untuk memilih data atau atribut yang akan digunakan dari database. Di tahapan seleksi inilah terjadinya pencarian data tersembunyi apabila data tidak relevan maka data akan dihapus. Dari hasil tahapan seleksi ini akan digunakan pada tahapan selanjutnya.

2. *Data Cleaning*

Dalam tahapan *cleaning* atau pembersihan data ini, bertugas untuk melakukan pembersihan data seperti menghapus data duplikasi, memeriksa data yang inkosisten, memperbaiki kesalahan pada data. Juga dilakukan proses memperkaya data yang sudah ada dengan data yang *relevan* atau *enrichment*.

3. *Data Transformation*

Pada tahapan *data transformation* data akan diubah maupun diinisialkan dalam bentuk *numeric*, *kategorikal*, *binominal* atau digabungkan kedalam format yang sesuai untuk proses data mining.

4. *Data Mining*

Data mining merupakan proses pencarian pola atau informasi dalam suatu data dengan menggunakan metode tertentu. Data mining berguna untuk menganalisa data pada data berskala besar sehingga menghasilkan suatu pola atau informasi yang bermanfaat.

5. *Interpretation/ Evaluation*

Interpretation/Evaluation merupakan suatu pola informasi yang dihasilkan dalam data mining untuk menampilkan suatu bentuk yang dapat mudah dimengerti oleh pihak yang berkepentingan. Tahapan ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

6. *Knowledge Presentation*

Pada *knowledge presentasion* ini akan menampilkan hasil dari penggunaan metode data mining sehingga pengguna dapat menemukan sebuah pengetahuan informasi baru.

2.4 Pengelompokkan *Data Mining*

Data mining memiliki beberapa kelompok yaitu *deskripsi*, *prediksi*, *estimasi*, *klasifikasi*, *clustering* dan *asosiasi*. Secara perinci proses data mining akan dijelaskan seperti yang dibawah ini. (Larose, 2005)

2.4.1 Deskripsi

Pada deskripsi ini bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli domain aplikasinya. Tugas deskripsi merupakan tugas data mining yang sering dibutuhkan teknik *postprocessing* untuk melakukan validasi dan menjelaskan hasil dari proses data mining. *Postprocessing* merupakan proses yang digunakan untuk memastikan hanya hasil yang valid dan berguna yang dapat digunakan oleh pihak berkepentingan.

2.4.2 Prediksi

Prediksi memiliki kemiripan dengan klasifikasi tetapi di prediksi data diklasifikasi berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh penerapan proses prediksi yaitu prediksi harga saham dalam tiga bulan yang akan datang berdasarkan data data yang ada di bursa saham.

2.4.3 Estimasi

Pada estimasi sama dengan prediksi kecuali variable target estimasi lebih kearah numerik dari pada kearah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variable target sebagai nilai prediksi. Selanjutnya pada peninjauan berikutnya estimasi nilai dari variable target dibuat berdasarkan nilai variable prediksi. Contohnya seperti hubungan antara tekanan darah sistolik dan nilai variable prediksi dalam proses pembelajaran akan menghasilkan model estimasi.

2.4.4 Klasifikasi

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data kedalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek kedalam salah satu kelas yang sudah didefinisikan sebelumnya.

2.4.5 Clustering

Clustering merupakan pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Dalam sebuah kluster berisi kumpulan

record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan record dalam kluster lain. *Clustering* memiliki tujuan yaitu untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok.

2.4.6 Asosiasi

Tugas *asosiasi* dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Tugas *asosiasi* berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

2.5 Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang akan terjadi dimasa depan berdasarkan informasi masa lalu dan informasi masa sekarang. Dalam prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, tetapi berusaha untuk mencari jawaban sedekat mungkin dengan yang akan terjadi di masa depan. Maka dapat disimpulkan secara sederhana prediksi sama dengan peramalan yang akan meramal atau memperkirakan sesuatu yang akan terjadi di masa yang akan datang. Dengan menggunakan data-data yang telah ada sebelumnya kita dapat meramalkan atau memprediksikan apa yang akan terjadi kedepannya (Herdianto, 2013).

2.6 Algoritma *K-Nearest Neighbor* (K-NN)

Algoritma *K-Nearest Neighbor* (K-NN) merupakan metode yang digunakan untuk mengelompokkan objek berdasarkan contoh latih terdekat di ruang fitur. Algoritma *K-Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama berdasarkan kecocokan bobot dari sejumlah fitur yang ada yang memiliki kesamaan. Algoritma *K-Nearest Neighbor* memiliki tujuan untuk mengklasifikasikan objek baru berdasarkan atribut dan training sample (Kusrini dan Emma, 2009).

Algoritma *K-Nearest Neighbor* akan mencari jarak terpendek dari sample uji (testing sample) ke sample latih (training sample) untuk menentukan *K-Nearest Neighbor* nya. Salah satu untuk menghitung jarak terdekat atau jauhnya tetangganya maka digunakan metode *euclidian distance*. Pada *euclidian distance*

sendiri berfungsi untuk menguji ukuran yang bisa digunakan sebagai interpretasi kedekatan jarak antara dua obyek, rumus *euclidian distance* seperti dibawah ini.

$$d(X,Y) = \sqrt{\sum_i^n (a_1 - b_1)^2 + (a_2 - b_2)^2, \dots, (a_n - b_n)^2} \dots\dots\dots (2.1)$$

Dimana a = a1, a2,an dan b = b1, b2bn mewakili n nilai atribut dari dua record.

Beberapa langkah-langkah untuk menghitung algoritma K-Nearest Neighbor yaitu sebagai berikut:

1. Pertama tentukan parameter K (jumlah tetangga terdekat)
2. Menghitung kuadrat jarak *Euclidian (queri instance)* masing-masing objek terhadap data sampel yang diberikan.
3. Kemudian mengurutkan jarak tersebut ke dalam kelompok yang memiliki *Euclid* terkecil.
4. Mengumpulkan kategori y berdasarkan nilai K atau dengan cara mengambil data tetangga terdekat.
5. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka akan menghasilkan kelas data baru.

2.7 Evaluasi

Evaluasi merupakan kunci yang ada pada pembuatan aplikasi berbasis *data mining*. Kinerja dari model kasifikasi dapat diukur dengan tingkat akurasi. Akurasi dari sebuah klasifikasi memberikan hasil latih dengan bentuk persentase dari kelompok data latih yang diklasifikasikan dengan benar (Han dkk, 2012). Cara mengetahui akurasi secara akurat dan untuk mengetahui seberapa baik prediksi, maka digunakanlah *confusion matrix*. TP dan TN akan memberikan informasi ketika prediksi benar. Sedangkan FP dan FN memberikan informasi katika prediksi salah. Dapat dilihat pada Tabel 2.1 di bawah ini.

Tabel 2.1 Akurasi *Confusion Matrix* (Han dkk, 2012)

No	Actual Class	Predicted Class	
		Positive	Negative
1.	Positive	TP	FN
2.	Negative	FP	TN

Akurasi merupakan persentase dari data yang diprediksi secara benar. Perhitungan akurasi adalah

$$\text{Akurasi} = \frac{(\text{TP}+\text{TN})}{(\text{TP}+\text{TN}+\text{FP}+\text{FN})} \dots\dots\dots (2.2)$$

Keterangan

TP : *True Positive*, merupakan banyaknya data yang kelas aktualnya adalah kelas positif dengan kelas prediksinya merupakan kelas positif.

TN : *True Negative*, merupakan banyaknya data yang kelas aktualnya adalah kelas positif dengan kelas prediksinya merupakan kelas negatif.

FP : *False Positive*, merupakan banyaknya data yang kelas aktualnya adalah kelas negatif dengan kelas prediksinya merupakan kelas positif.

FN : *False Negative*, merupakan banyaknya data yang kelas aktualnya negatif dengan kelas prediksinya merupakan kelas negatif.

2.8 Penelitian Terkait

Berikut beberapa penelitian terkait dengan penelitian berkaitan penyakit mata pad Tabel 2.2 di bawah ini:

Tabel 2.2 Penelitian Terkait

No.	Peneliti dan Tahun	Judul	Kesimpulan
1.	Novita Mariana Tahun 2015 ISSN 2085-3343	Penerapan Algoritma KNN (K Nearest Neighbor) Untuk Deteksi Penyakit Kanker Serviks	Tujuan penelitian ini untuk dapat membantu pengguna mencari informasi mengenai penyakit kanker serviks dan konsultasi keluhan gejala penyakit kanker serviks dengan mudah.
2.	Rizal Yepriyanto Tahun 2017 ISSN 1693-1173	Sistem Diagnosa Kesuburan Sperma Dengan Metode K-Nearest Neighbor (KNN)	Tujuan penelitian ini untuk membuat system diagnosa kesuburan sperma dengan menggunakan metode KNN berdasarkan 20 data uji dan 17 data klasifikasi system dapat memberikan akurasi system 85%
3.	Olha Musa Tahun 2017 ISSN 2548-7779	Analisa Penyakit Paru-Paru Menggunakan Algoritma KNN pada Rumah Sakit Aloj Saboe Kota Gorontalo	Penelitian ini bertujuan untuk melihat akurasi system dalam penggunaan metode KNN dengan medeteksi penyakit paru paru. Akurasi yang didapat dari system ini 91,90%

No.	Peneliti dan Tahun	Judul	Kesimpulan
4.	Wiyli yustanti Tahun 2012 ISSBN:1858-1382	Algoritma K-Nearst Neighbour untuk memprediksi harga jual tanah	Tingkat presisi dari prediksi yang dihasilkan dalam jurnal ini sebesar 80 % untuk meningkatkan akurasinya maka diperlukan tambahan data sample
5.	Yerry anggoto dkk Tahun 2018	Implementasi metode fuzzy K-nearest neighbor untuk klasifikasi penyakit tanaman kedelai pada citra daun	Akurasi yang didapat pada penelitian ini sebesar 83,3 persen. Pada pengujian lain menggunakan nilai k didapat akurasi 83,3 %
6.	Ari Muzakir ISSBN 2460-0040 Tahun 2016	Model data mining sebagai prediksi penyakit hipertensi kehamilan dengan teknik decision tree	Tingkat akurasi yang didapat dalam penelitian ini 92,6573 %. Penelitian ini menghasilkan informasi berupa prediksi penyakit hipertensi dalam kehamilan.
7.	Henny Leidiyana Tahun 2013	Penerapan Algoritma K-Nearest Neighborn untuk penentuan resiko kredit kepemilikan kendaraan bermotor	Hasil testing untuk mengukur performa algoritma ini menggunakan metode Cross Validation, Confusion Matrix dan kurvaROC dan menghasilkan akurasi dan nilai AUC berturut-turut 81,46 % dan 0,984. Karena nilai AUC berada dalam rentang 0,9 sampai 1,0 maka metode tersebut masuk dalam kategori sangat baik (excellent)
8.	Mutiara Ayu Banjarsari1, H. Irwan Budiman, Andi Farmadi3 ISSN: 2406-7857 Tahun 2015	Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4	Kesimpulan yang diperoleh dari penelitian ini adalah: a. Nilai k-Optimal pada algoritma kNN untuk prediksi kelulusan tepat waktu mahasiswa berdasarkan IP sampai dengan semester 4 adalah k=5. b. Dari proses k-Fold Cross Validation didapatkan tingkat akurasi untuk k=5 pada algoritma kNN untuk prediksi kelulusan tepat waktu mahasiswa berdasarkan IP sampai dengan semester 4 adalah sebesar 80,00%.

No.	Peneliti dan Tahun	Judul	Kesimpulan
9.	Nobertus risandi, Helmi, Bayu Prihandono Tahun 2013	Algoritma K-nearest Neighbor dalam Klasifikasi Data Hasil Produksi Kelapa Sawit Pada PT Minamas Kecamatan Parindu	Berdasarkan hasil penelitian data terklasifikasi dalam 6 cluster berdasarkan kemiripan hasil produksi dari 50 kelompok tani yang ada di KUD. HIMADO. Hasil produksi yang dominan adalah produksi dari kelompok tani kelapa sawit yang terletak pada C1 dengan 17 anggota dengan persentase 34% yaitu kelompok 1, 2, 33, 34, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49,50 untuk nilai k=7.
10.	Muhammad Rivki dan Adam Mukharil Bachtiar Tahun 2017	Implementasi Algoritma K-nearest Neighbor dalam Pengklasifikasian Follower Twitter yang Menggunakan Bahasa Indonesia	Berdasarkan hasil implementasi dan pengujian pada aplikasi Twikipedia maka diperoleh kesimpulan sebagai berikut: 1) Aplikasi ini dapat membantu memudahkan pengguna Twitter yang menggunakannya sebagai media pemasaran atau promosi untuk melakukan promosi dengan cara melakukan tweet promosi terhadap follower yang sudah diklasifikasikan. 2) Aplikasi ini dapat membantu memudahkan pengguna Twitter yang menggunakannya setabbagai media pemasaran atau promosi untuk menentukan waktu tweet pormosi yang tepat yang sudah dianalisa dari waktu aktif follower pengguna.