

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

1.1 Tinjauan Pustaka

Ida Bagus Adisimakrisna Peling, I Nyoman Arnawan, I putu Arich Arthawan, dan I Gusti Ngurah Janardana (2017), melakukan penelitian kualitas perguruan tinggi khususnya program studi di Indonesia diukur berdasarkan akreditasi yang dilakukan oleh BAN PT. Menurut BAN PT kualitas diukur berdasarkan 7 standar utama, salah satunya mahasiswa dan pascasarjana. Salah satu permasalahan yang masih menjadi bahan diskusi terkait kegagalan mahasiswa adalah mahasiswa yang tidak tepat waktu. Mahasiswa yang lulus tidak tepat waktu adalah mahasiswa yang tidak dapat menyelesaikan studinya sesuai dengan ketentuan waktu yang diberikan. Keberadaan mahasiswa pascasarjana yang tidak tepat waktu tentunya menimbulkan permasalahan dan berpotensi putus sekolah yang berdampak pada kualitas pendidikan dan akreditasi. Diperlukan sistem yang memprediksi kelulusan siswa dengan mengevaluasi hasil belajarnya. Ketepatan waktu kelulusan mahasiswa dapat dilakukan dengan teknik data mining untuk mengetahui pola kelulusan mahasiswa yang telah lulus yang kemudian dijadikan dasar untuk memprediksi kelulusan mahasiswa tahun depan.

Hasil penelitian menunjukkan bahwa *Naïve Bayes Classification (NBC)* mampu mengklasifikasikan pengujian data yang benar secara rata-rata sebesar 86.16% dan error 13.84%. Selain itu, informasi lain yang diperoleh dari data pengujian yang digunakan yaitu siswa yang masuk dari PMDK Lulus lulus tepat waktu sebanyak 40%, jalur lain lulus tepat waktu sebesar 26,7%, dan lulus ujian filter tepat waktu 13,3%.

Muhammad Ibnu Hawari, Bosker Sinaga (2019) melakukan penelitian Lateks adalah produksi yang dihasilkan dari pohon karet dengan cara disadap menggunakan pisau khusus, dan di dalam perkebunan terdapat beberapa hasil tanaman yang diolah di dalam tanaman itu sendiri menjadi bahan dasar atau barang jadi bagi konsumen.

Selama ini penentuan prediksi hasil produksi lateks tidak menggunakan metode sehingga hasil tidak sesuai dengan apa yang diharapkan. Hal ini disebabkan belum adanya metode yang obyektif untuk menentukan prediksi hasil produksi lateks, pemilihan yang cepat berdasarkan data pekerja penyadap beberapa pohon karet dengan beberapa teknik dan cara untuk mendapatkan hasil yang baik. Dengan mengacu pada solusi yang diberikan algoritma *Naïve Bayes Classification (NBC)* untuk membantu memprediksi hasil produksi lateks, Seorang leader mengumpulkan hasil getah dari 3 item yaitu Latex, Lump dan Treelace.

Perhitungan dilakukan dengan metode algoritma item ketiga Naïve Bayes Classification (NBC) dengan klasifikasi yang menerapkan data mining dan data probabilitas. Data Mining umumnya didefinisikan sebagai sistem yang mampu menghasilkan dan menangani pemecahan masalah. Algoritma Naïve Bayes Classification (NBC) dapat menentukan nilai preferensi masing-masing alternatif, dan dapat menjadi alternatif terbaik dari sejumlah alternatif.

Edi Sutoyo dan Ahmad Almaarif (2020) melakukan penelitian kualitas mahasiswa dapat dilihat dari prestasi akademik yang merupakan bukti dari upaya yang dilakukan oleh mahasiswa. Prestasi akademik siswa dievaluasi setiap akhir semester untuk mengetahui hasil belajar yang telah dicapai. Jika seorang mahasiswa tidak dapat memenuhi kriteria akademik tertentu yang ditetapkan dengan memenuhi persyaratan untuk melanjutkan studinya, maka mahasiswa tersebut berpotensi untuk tidak lulus tepat waktu atau bahkan Drop Out (DO). Tingginya jumlah mahasiswa yang tidak lulus tepat waktu atau DO di perguruan tinggi dapat diminimalisir dengan mendeteksi mahasiswa yang berisiko pada tahap awal pendidikan dan didukung dengan pembuatan kebijakan yang dapat mengarahkan mahasiswa untuk menyelesaikan pendidikannya. Selain itu, jika waktu penyelesaian studi siswa dapat diprediksi maka penanganan siswa akan lebih efektif.

Salah satu teknik untuk membuat prediksi yang dapat digunakan adalah teknik data mining. Oleh karena itu, dalam penelitian ini algoritma Naïve Bayes Classification (NBC) akan digunakan untuk memprediksi kelulusan mahasiswa

Telkom University. Dataset diperoleh dari Direktorat Sistem Informasi (SISFO) Telkom University yang berisi 4000 data instans. Hasil penelitian ini membuktikan bahwa NBC berhasil diterapkan untuk memprediksi kelulusan mahasiswa.

Prediksi kelulusan mahasiswa tersebut mampu menghasilkan akurasi 73,725%, presisi 0,742, recall 0,736 dan F-measure 0,735. Dataset diperoleh dari Direktorat Sistem Informasi (SISFO) Telkom University yang berisi 4000 data instans. Hasil penelitian ini membuktikan bahwa NBC berhasil diterapkan untuk memprediksi kelulusan mahasiswa. Prediksi kelulusan mahasiswa tersebut mampu menghasilkan akurasi 73,725%, presisi 0,742, recall 0,736 dan F-measure 0,735. Dataset diperoleh dari Direktorat Sistem Informasi (SISFO) Telkom University yang berisi 4000 data instans. Hasil penelitian ini membuktikan bahwa NBC berhasil diterapkan untuk memprediksi kelulusan mahasiswa. Prediksi kelulusan mahasiswa tersebut mampu menghasilkan akurasi 73,725%, presisi 0,742, recall 0,736 dan F-measure 0,735.

Puji Hari Santoso, Fauziah Fauziah, Nurhayati Nurhayati (2020) melakukan penelitian status terjangkit virus yang dikenal dengan Covid -19 semakin meningkat di wilayah Jakarta Selatan yaitu Pondok Labu, Cilandak Barat, Jagakarsa, Lenteng Agung, Pasar Minggu dan Ragunan, maka perlu dilakukan pengklasifikasian data untuk mengetahui status negatif penyakit tersebut. infeksi virus covid-19 atau positif terinfeksi virus covid -19.

Teknik pengklasifikasian status virus Covid-Virus positif atau negatif dengan metode klasifikasi naïve bayes. Untuk mengelola data digunakan software rapid miner 9. 6, dataset status terinfeksi covid -19 diperoleh dari website [jeo.kompas](#). Perhitungan Prediksi menunjukkan klasifikasi metode Naïve Bayes diperoleh prediksi positif yang menunjukkan angka 55,48% dan hasil prediksi negatif 44,52%.

Bustami Yusuf, Muthmainna Qalbi, Basrul Basrul, Ima Dwitawati, Malahayati Malahayati, Mega Ellyadi (2020) melakukan penelitian Prestasi akademik ditentukan oleh dua faktor yaitu faktor internal yang berasal dari dalam diri individu dalam hal ini mahasiswa dan faktor eksternal yang berasal dari luar individu atau hal-hal yang dipengaruhi oleh lingkungan. Ada banyak cara untuk menemukan suatu prestasi akademik, salah satunya dengan menggunakan data mining yang bertujuan untuk memprediksi atau mengklasifikasikan data menggunakan algoritma klasifikasi. Penelitian ini bertujuan untuk 1) mengetahui bagaimana penerapan algoritma Naive Bayes terhadap prestasi belajar siswa, dan 2) melihat keakuratan algoritma Naive Bayes terhadap prestasi belajar siswa. Jenis penelitian ini adalah data sekunder berupa data mahasiswa yang diperoleh dari pusat teknologi informasi dan basis data UIN Ar-Raniry. Penelitian ini menggunakan algoritma Naive Bayes dan algoritma random forest.

Hasil yang diperoleh dari penelitian ini menunjukkan nilai korelasi tertinggi pada variabel IP awal $r = 0,783$ dan variabel cuti memiliki tingkat korelasi yang sangat lemah yaitu $r = 0,054$. Nilai akurasi algoritma *Naïve Bayes Classification (NBC)* setelah pembersihan sebesar 78.0% dan variabel algoritma Random Forest sebesar 76.7%.

Tabel 2.1 Perbandingan Penelitian Yang Sudah Ada.

No	Nama Pengarang	Judul	Objek	Metode	Hasil
1	Ida Bagus Adisimakrisna Peling, I Nyoman Arnawan, Iputu Arich Arthawan, dan I Gusti Ngurah Janardana (2017)	Implementasi Data Mining Untuk Memprediksi Masa Belajar Siswa Menggunakan Algoritma Naive Bayes	Data kualitas perguruan tinggi berdasarkan akreditasi yang dilakukan oleh BAN PT.	Naïve bayes	pengujian data yang benar secara rata-rata sebesar 86.16% dan error 13.84%.
2	Muhammad Ibnu Hawari, Bosker Sinaga (2019)	Implementasi Algoritma Naïve Bayes Untuk Memprediksi Produksi Gum di PT. Pengadilan Sri Rahayu	Data penelitian Lateks adalah produksi yang dihasilkan dari pohon karet.	Naïve bayes	hasil getah dari 3 item yaitu Latex, Lump dan Treelace.
3	Edi Sutoyo dan Ahmad Almaarif (2020)	Penambangan Data Pendidikan Untuk memprediksi Kelulusan Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier	Data mahasiswa berpotensi untuk tidak lulus tepat waktu atau bahkan Drop Out (DO).	Naïve bayes	menghasilkan akurasi 73,725%, presisi 0,742, recall 0,736 dan F-measure 0,735.

4	Puji Hari Santoso, Fauziah Fauziah, Nurhayati Nurhayati (2020)	Penerapan Klasifikasi Data Mining untuk Status Terinfeksi Covid-19 Menggunakan Metode Algoritma Naïve.	Data status negatif penyakit Covid- 19 tersebut, di wilayah Jakarta Selatan yaitu Pondok Labu.	Naïve Bayes	prediksi positif yang menunjukkan angka 55,48% dan hasil prediksi negatif 44,52%.
5	Yusuf Bustami, Qalbi Muthmainna, Basrul, Dwitawati Ima, Malahayati, Ellyadi Mega (2020)	Implementasi Algoritma Naïve Bayes dan Random Forest Dalam Memprediksi Prestasi Akademik Mahasiswa Universitas Islam Negeri Ar- Raniy Banda Aceh.	Data mahasiswa dari pusat teknologi informasi dan basis data UIN Ar-Raniy	Naïve Bayes, Random Forest	Menggunakan nilai toleransi tinggi dan nilai akurasi algoritma naïve bayes.

Pada tabel 2.1 ini adalah hasil perbandingan dari penelitian sebelumnya.

1.2 Dasar Teori

Dasar teori ini untuk memahami definisi, pengertian dasar dan istilah yang digunakan untuk penelitian ini. Berikut dasar teori yang digunakan.

1.2.1 Sistem Penilaian Prestasi Mahasiswa

Mahasiswa adalah bagian dari sivitas akademika yang merupakan peserta didik berstatus aktif pada semester yang berjalan pada salah satu program studi yang diselenggarakan oleh Sekolah Tinggi.

Kegiatan ekstrakurikuler adalah kegiatan kemahasiswaan di luar kegiatan akademik yang meliputi pengembangan penalaran dan keilmuan, minat dan bakat, kewirausahaan, pengabdian masyarakat, upaya perbaikan kesejahteraan mahasiswa, dan pengembangan organisasi kemahasiswaan, yang dapat dilakukan di dalam maupun di luar kampus.

Sistem Penilaian Prestasi Mahasiswa (SPPM) adalah mekanisme dan aturan yang digunakan untuk menyatakan pengakuan atas prestasi pengembangan *softskill* kemahasiswaan.

Guna mendorong mahasiswa melakukan kegiatan dan berorganisasi sehingga akan meningkatkan rasa percaya diri dalam mengikuti kompetisi. Sistem ini merupakan kebijakan dan mekanisme mengenai penghargaan kegiatan kemahasiswaan khususnya yang bersifat ekstrakurikuler.

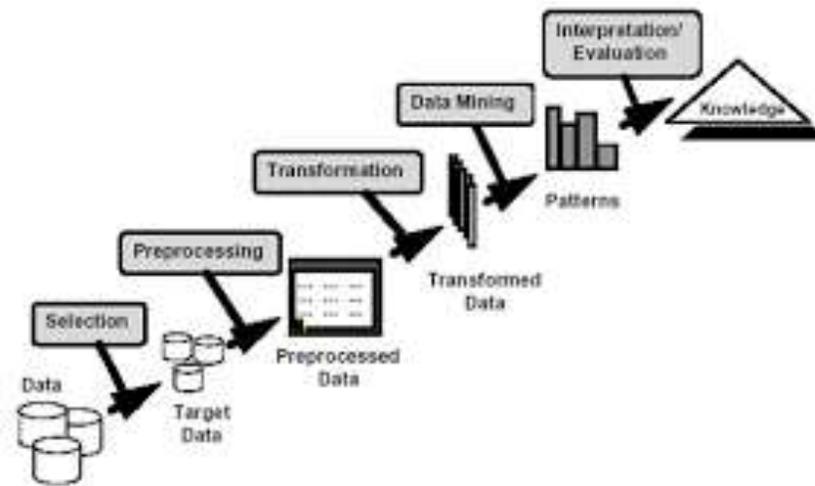
1.2.2 Data Mining

Menurut Tan dalam Prasetyo (2012:2) “Data Mining adalah suatu proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan”.

Tahap-tahap data mining Menurut Fayyad (1996), istilah data mining dan *Knowledge Discovery in Basis datas* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam proses KDD adalah data mining. Proses KDD secara garis besar dapat dijelaskan sebagai berikut:

- a. Data Pemilihan seleksi data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining. Disimpan dalam suatu berkas, terpisah dari basis data operasional.
- b. Pembersihan Data sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses pembersihan data pada data yang menjadi fokus KDD. Proses pembersihan data mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.
- c. Transformasi adalah proses transformasi pada data yang telah dipilih, sehingga tersebut sesuai untuk proses data mining. Proses coding dalam data KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.
- d. Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

- e. Evaluasi pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut evaluasi.



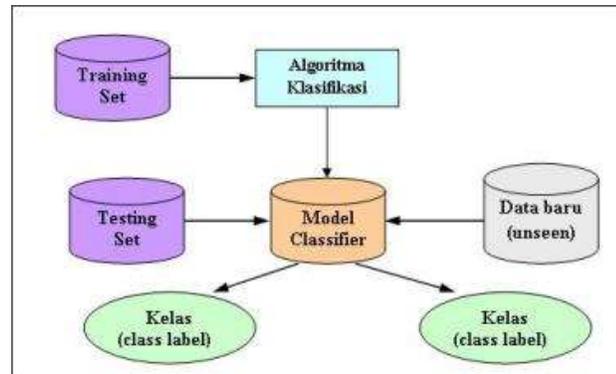
Gambar 2.1 Proses KDD

Pada gambar 2.1 adalah proses KDD untuk pembersihan data dengan beberapa tahapan.

1.2.3 Klasifikasi

Salah satu pengukur kinerja klasifikasi adalah tingkat akurasi. Sebuah sistem dalam melakukan klasifikasi diharapkan dapat mengklasifikasi semua set data dengan benar, tetapi tidak dipungkiri bahwa kinerja suatu sistem tidak bisa 100% akurat. (Prasetyo, 2012). Pada 2.1 akurasi untuk menghitung akurasi digunakan rumus :

$$\text{Akurasi} = \frac{\text{jumlah data yang diprediksi secara benar}}{\text{jumlah prediksi yang dilakukan}} \quad (2.1)$$



Gambar 2.2 Proses pengerjaan klasifikasi

Pada gambar 2.2 ini adalah proses klasifikasi untuk menghitung semua data apakah data benar dan diujikan dengan hitung akurasi.

1.2.4 Algoritma *Naïve Bayes Classification (NBC)*

Naïve bayes merupakan teknik prediksi berbasis probabilitik sederhana yang berdasar pada penerapan teorema Bayes (aturan Bayes) dengan asumsi independensi (ketidaktergantungan) yang kuat (naif). Dengan kata lain, dalam *Naïve Bayes Classification (NBC)* model yang digunakan adalah “model fitur independen”. Fitur independensi yang kuat adalah sebuah fitur pada sebuah data tidak ada kaitannya dengan adanya atau tidak adanya fitur yang lain dalam data yang sama atau dengan kata lain *Naïve Bayes Classification (NBC)* menganggap bahwa efek dari nilai atribut pada kelas tertentu tidak bergantung pada nilai atribut lainnya. (Han & Kamber, 2012).

Pendekatan ini didasarkan pada kuantifikasi trade-off antara berbagai keputusan klasifikasi dengan menggunakan probabilitas dan ongkos yang ditimbulkan dalam keputusan-keputusan tersebut. Ide dasar dari bayes adalah menangani masalah yang bersifat hipotesis yakni mendesain suatu klasifikasi untuk memisahkan objek (Amalia, N, dkk. 2017).

Bayesian classification adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Bayesian classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam *basis data* dengan data yang besar (Kusrini dan Luthfi ,2009). Prediksi bayes didasarkan pada *Naïve Bayes Classification (NBC)* dengan bentuk umum seperti pada rumus Naïve Bayes pada rumus 2.2:

$$P(C_i | X) = p(X | C_i)P(C_i) p(X) \dots \quad (2.2)$$

Keterangan :

- X : Data dengan Class yang belum diketahui
- C_i : Suatu variabel yang harus dideskripsikan secara probabilistik
- P(C_i|X): Probabilitas hipotesis C_i berdasarkan kondisi X (posteriori probability)
- P(C_i): Probabilitas hipotesis C_i (prior probability)
- P(X|C_i): Probabilitas X berdasar kondisi pada hipotesis C_i

- P(X): Probabilitas dari X Berikut ini merupakan contoh data training sebanyak 14 data dengan output main sepak bola atau tidak. Setiap data ditandai dengan atribut cuaca, temperatur, kelembaban, dan angin yang dapat dilihat pada Tabel 2.2.

Tabel 2.2 Data Training Cuaca dan Keputusan Main atau Tidak

Cuaca X1	Temperatur X2	Kelembaban X3	Angin X4	Main Atau Tidak Y
Cerah	Panas	Tinggi	Kecil	Tidak
Cerah	Panas	Tinggi	Besar	Tidak
Mendung	Panas	Tinggi	Kecil	Ya
Hujan	Sedang	Tinggi	Kecil	Ya
Hujan	Dingin	Normal	Kecil	Ya
Hujan	Dingin	Normal	Besar	Tidak
Mendung	Dingin	Normal	Besar	Ya
Cerah	Sedang	Tinggi	Kecil	Tidak
Cerah	Dingin	Normal	Kecil	Ya
Hujan	Sedang	Normal	Kecil	Ya
Cerah	Sedang	Normal	Besar	Ya
Mendung	Sedang	Tinggi	Besar	Ya
Mendung	Panas	Normal	Kecil	Ya
Hujan	Sedang	Tinggi	Kecil	Tidak

Pada tabel 2.2 berikut ini merupakan contoh perhitungan metode Naïve Bayes yang menggunakan rumus Naïve Bayes (2.2) untuk menentukan kelas dari data baru berikut:

(Cuaca= cerah, Temperatur= dingin, Kelembaban= tinggi, Angin= besar) Langkah – langkah perhitungan yang dikerjakan sebagai berikut:

a. Mencari probabilitas prior

Probabilitas prior ini menyatakan berapa peluang munculnya keputusan main sepak bola $P(C1)$ dan peluang keputusan tidak main sepak bola $P(C2)$. Misalkan N adalah jumlah total keputusan main sepak bola dan tidak, $N1$ menyatakan jumlah keputusan main sepak bola dan $N2$ menyatakan jumlah keputusan tidak main sepak bola. Berdasarkan pada Tabel 2.2 jumlah total keputusan main sepak bola dan tidak sebanyak 14, jumlah keputusan main sepak bola sebanyak 9, dan jumlah keputusan tidak main sepak bola sebanyak 5. Didapatkan rumus sebagai berikut :

$$P(C1) = \frac{N1}{N} = \frac{9}{14} = 0,64$$

$$P(C2) = \frac{N2}{N} = \frac{5}{14} = 0,36$$

b. Mencari Probabilitas bersyarat (likelihoad)

Probabilitas bersyarat (likelihoad), $P(X|Ci)$ ini menyatakan peluang munculnya X jika diketahui Ci . Misalkan :

$X:(Cuaca=cerah, Temperatur=dingin, Kelembaban=tinggi, Angin=besar)$

Ci : $C1$ adalah main sepak bola dan $C2$ adalah tidak main sepak bola.

Sehingga likelihoad dapat dihitung dengan mengalikan hasil dari masing – masing nilai probabilitas per-atribut yang berdasarkan pada Ci .

Langkah – langkah menghitung likelihood :

- Menghitung probabilitas per-atribut

Misalkan menghitung probabilitas Angin = besar dengan Ci adalah C1 yaitu main sepak bola. Berdasarkan Tabel 2.2 banyaknya Angin = besar yang keputusannya adalah main sepak bola ada 3 dan banyaknya kelas dengan keputusan main sepak bola ada 9. Maka probabilitasnya adalah

$$(Angin = Besar | main) = \frac{\text{jumlah Angin Besar dengan keputusan main}}{\text{jumlah seluruh keputusan main}}$$

$$\text{atau} = \frac{3}{9} = 0,33.$$

Melakukan perhitungan dengan cara yang sama pada probabilitas per-atribut yang lain, sehingga diperoleh hasil :

$$P(\text{Cuaca=Cerah}|\text{Main}) = \frac{2}{9} = 0,22$$

$$P(\text{Cuaca=Cerah}|\text{Tidak}) = \frac{3}{5} = 0,60$$

$$P(\text{Temperatur=dingin} | \text{Main}) = \frac{3}{9} = 0,33$$

$$P(\text{Temperatur=dingin} | \text{Tidak}) = \frac{1}{5} = 0,20$$

$$P(\text{kelembaban=tinggi} | \text{Main}) = \frac{3}{9} = 0,33$$

$$P(\text{kelembaban=tinggi} | \text{tidak}) = \frac{4}{5} = 0,80$$

$$P(\text{Angin=Besar} | \text{Tidak}) = \frac{3}{5} = 0,60$$

- Mengalikan semua probabilitas per-atribut yang berdasarkan pada Ci yaitu C1 adalah main yang disebut Likelihood Ya atau C2 adalah tidak main yang disebut Likelihood Tidak.

- o Likelihood Ya

$$P(X|Main) = P(\text{Cuaca=Cerah}|Main) * P(\text{Temperatur=dingin} | Main) * P(\text{kelembaban=tinggi} |Main) * P(\text{Angin=Besar} |Main) \\ = 0,22 * 0,33 * 0,33 * 0,33 = 0.0080$$

- o Likelihood Tidak

$$P(X|Tidak) = P(\text{Cuaca=Cerah}|Tidak) * P(\text{Temperatur=dingin} | Tidak) * P(\text{kelembaban=tinggi} |Tidak) * P(\text{Angin=Besar} |Tidak) \\ = 0.60 * 0,20 * 0,80 * 0,60 = 0.0576$$

c. Mengalikan Likelihood dengan Prior

Perkalian Likelihood dengan Prior merupakan hal yang penting untuk menemukan posterior.

- o Perkalian Likelihood dengan Prior pada keputusan Main

$$= P(X|Ci) * P(Ci) = P(X|Main) * P(Main) = 0.0080 * 0,64 = 0.00512$$

- o Perkalian Likelihood dengan Prior Tidak Main

$$= P(X|Ci) * P(Ci) = P(X|Tidak) * P(Tidak) = 0.0576 * 0,36 = 0.020736$$

d. Menghitung probabilitas posterior

Probabilitas posterior, $P(C_i|X)$ ini menyatakan probabilitas keluarnya hasil C_i jika diketahui nilai X tertentu. Probabilitas posterior, $P(C_i|X)$ dicari dengan menggunakan rumus pada (2.2) :

- Probabilitas Posterior Main

$$P(C_i | X) = \frac{p(X | C_i)P(C_i)}{p(X)}$$

$$posterior = \frac{likelihood * prior}{evidenc}$$

$$= \frac{p(X | Main) * P(Main)}{P(X | Main) * P(Main) + P(X | Tidak) * P(Tidak)} = \frac{0.00512}{0.00512 + 0.020736}$$

$$= 0,198019802$$

- Probabilitas Posterior Tidak Main

$$P(C_i | X) = \frac{p(X | C_i)P(C_i)}{p(X)}$$

$$posterior = \frac{likelihood * prior}{evidence}$$

$$= \frac{P(X | Tidak) * P(Tidak)}{P(X | Tidak) * P(Tidak) + p(X | Main) * P(Main)} = \frac{0.020736}{0.020736 + 0.00512}$$

$$= 0,801980198$$

Nilai probabilitas posterior Tidak Main sebesar 0,801980198 lebih besar dari nilai probabilitas posterior Main yang bernilai 0,198019802 dan nilai probabilitas posterior Tidak Main mendekati nilai 1, sehingga dengan *Naïve Bayes Classification (NBC)* prediksi dapat disimpulkan Tidak Main untuk data input ini yang berdasarkan pada estimasi probabilitas yang dipelajari dari data *training*.

1.2.5 MySQL

MySQL adalah salah satu aplikasi pengolahan basis data yang sering digunakan oleh banyak programmer, disamping sudah mendukung dalam pembuatan basis data yang berbasis Client atau Server, juga dapat mengolah basis data dalam jumlah besar. SQL adalah bahasa standar untuk query yang digunakan untuk manipulasi data yang dibuat berbagai DBMS (*Basis data, Management, System*) salah satunya adalah MySQL.

Beberapa keistimewaan yang dimiliki oleh MySQL adalah *Probability, Open Source, Multiuser, Performer Tuning, Column Tuning, Command and Function, Security, Scalability and Limits, Connectivity, Localization, Interface, Client and Tools, Structure Tabel* (S. Mutiah, 2019).

1.2.6 PHP

PHP menjadikan website yang dibangun mampu mengolah data yang dimasukkan pengguna dan menghasilkan luaran yang diinginkan. Penelitian yang telah dilakukan menunjukkan bahwa PHP dapat digunakan sebagai pendukung HTML guna menghasilkan website yang mampu mengolah data (Janpla, Kularbphetong, Rattanachai, & Tangtastham, 2015; Surbakti & Pasaribu, 2018).