

BAB II

TINJAUAN PUSTAKA

1.1 Penelitian Terdahulu

Dalam penelitian Naparin (2016) yang berjudul Klasifikasi Peminatan Siswa SMA Menggunakan Metode Naive Bayes. Penelitian ini menjelaskan bahwa prinsip Bayes atau prinsip Probabilitas yaitu prinsip yang berdasarkan hasil observasi dan difokuskan pada proses pengklasifikasian menggunakan metode Naive Bayes yang memegang asumsi akan hubungan antar fitur atau atributnya yang independen sehingga menjadikannya lebih efektif untuk kategorisasi, sederhana, cepat dan menghasilkan tingkat akurasi yang tinggi.

Dalam penelitian Soemantri (2017) yang berjudul Text Mining Untuk Klasifikasi Kategori Cerita Pendek Menggunakan Naïve Bayes (NB). Penelitian ini menjelaskan tentang penerapan metode Naïve Bayes untuk pengklasifikasian jenis kategori cerita pendek sehingga didapatkan tingkat akurasi yang terbaik untuk klasifikasi kategori cerita pendek.

Dalam penelitian Harimurti (2017) yang berjudul Klasifikasi Penerimaan Beasiswa Menggunakan Metode Naïve Bayes Classifier (Studi Kasus Universitas Trunojoyo Mandura). Penelitian ini menjelaskan bahwa studi kasus yang menjadi latar penelitian selama ini untuk menyeleksi penerimaan beasiswa masih menggunakan Excel atau secara manual karena proses seleksi yang dilakukan akan memakan waktu yang lama dan kurang efektif. Sehingga perlu adanya kecanggihan teknologi untuk mengatasi suatu pekerjaan tersebut, maka dengan menerapkan metode Naïve Bayes Classifier diharapkan dapat digunakan dalam hal pengambilan keputusan untuk mendapatkan hasil yang lebih baik pada suatu permasalahan klasifikasi.

Dalam penelitian Faridah, Ulinuha, (2018) yang berjudul Klasifikasi Mahasiswa Penerima Program Beasiswa Bidik Misi Menggunakan Naïve Bayes. Penelitian ini menjelaskan bahwa studi kasus latar penelitian ini masih menggunakan cara manual dalam menyeleksi apalagi dengan melihat indikator dalam penyeleksian berkas pengajuan beasiswa dengan beberapa macam kriteria,

kelengkapan data yang dikumpulkan terkesan objektif, kurang efektif dan membutuhkan waktu relative lebih lama. Sehingga perlu adanya metode khusus dan sistem tertentu yang dapat memberikan suatu keputusan yang tepat, efektif dan efisien dalam proses penyeleksian untuk menentukan penerima beasiswa bidik misi bagi seluruh pendaftar berdasarkan data yang masuk.

Dalam penelitian Mauludy, (2020) yang berjudul Klasifikasi Rumpun Keilmuan di Perguruan Tinggi Keagamaan Islam Negeri menggunakan Metode Naïve Bayes Classifier. Penelitian ini akan mengklasifikasin jurusan-jurusan sesuai pada bidangnya masing-masing menggunakan silabus. Yang mana rumpun keilmuan ini dapat memudahkan para mahasiswa baru untuk memilih jurusan sesuai bakat dan minatnya.

Sehingga dapat diambil kesimpulan bahwa memberi keputusan dengan banyak pertimbangan akan menjadikan suatu permasalahan jika tidak dipertimbangkan secara teliti dan hati-hati. Maka untuk mepermudah pengambilan suatu keputusan yang baik dan akurat terkhusus dalam mengklasifikasi pembagian suatu kelas atau kelompok perlu adanya alat bantu teknologi yang menjadikan keputusan tersebut benar-benar dapat dipercaya.

1.2 Landasan Teori

1.2.1 Klasifikasi

Klasifikasi merupakan kata serapan dari Bahasa Belanda *classificatie*, yang sendirinya berasal dari Bahasa Prancis *Classification*. Istilah ini menunjukkan kepada sebuah metode untuk menyusun data secara sistematis atau menurut beberapa kaidah atau aturan yang telah ditetapkan.

Secara harfiah bisa pula dikatakan bahwa klasifikasi adalah pembagian sesuatu menurut kelas-kelas. Menurut Ilmu Pengetahuan, klasifikasi adalah Proses pengelompokan benda berdasarkan ciri-ciri perbedaan dan persamaan (Kamus Bahasa Indonesia).

Klasifikasi juga merupakan proses pembelajaran suatu fungsi tujuan (target) yang menetapkan tiap himpunan atribut x ke satu dari label kelas y yang di definisikan sebelumnya. Dalam mencapai tujuan tersebut, proses klasifikasi

membentuk suatu model yang mampu membedakan data kedalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu. Model itu sendiri bisa berupa aturan berupa pohon keputusan atau formula matematis. (Nurjaya, dkk.,2018)

Tujuan klasifikasi adalah untuk:

1. Menentukan model dari data latih yang membedakan record ke dalam atau kategori yang sesuai, model tersebut kemudian digunakan untuk mengklasifikasi record yang kelasnya belum diketahui sebelumnya pada testing set.
2. Mengambil keputusan dengan meprediksi suatu kasus, berdasarkan hasil klasifikasi yang diperoleh (Nugraha dkk.,2013).

1.2.2 Konsep Klasifikasi

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkan kedalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu: pertama, Pembangunan model sebagai prototype untuk disimpan sebagai memoru dan kedua, Penggunaan model tersebut untuk melakukan pengenalan/klasifikasi/prediksi pada suatu objek data lain agar diketahui di kelas mana objek tersebut dalam model yang mudah disimpan (Patil dkk., 2012)

Klasifikasi juga merupakan pengelompokan beberapa parameter kedalam salah satu kategori yang sudah ditetapkan sebelumnya. Klasifikasi sering disebut dengan supervised learning hal tersebut disebabkan karena kelas yang digunakan dalam proses pengklasifikasian telah ditentukan sebelumnya. Pada umumnya, klasifikasi dipakai untuk melakukan prediksi suatu label terhadap kelas tertentu. Hal tersebut dilakukan dengan membangun suatu model klasifikasi menggunakan sejumlah data training. Kemudian dari model klasifikasi yang dihasilkan, akan digunakan untuk mengklasifikasi data yang baru.

1.2.3 Santri Kelas Tahfidz

Menurut Kamus Besar Bahasa Indonesia, pengertian santri adalah orang yang mendalami agama islam, orang yang beribadat dengan sungguh-sungguh, orang sholeh. Sedangkan pengertian Tahfidz menurut Mahmud Yunus (1990:105) berarti

menghafal berasal dari kata dasar hafal yang bahasa arab *hafidzo – yahfadzu – hifdzan*, yaitu lawan dari lupa, yaitu sellau ingat dan sedikit lupa.

Tahfidz menurut Sumardi (1993:45) merupakan metode menghafal Al Qur'an yaitu aktivitas mencamkan dengan sengaja dan dikehendaki dengan sadar dan sungguh-sungguh. Begitu juga dengan menurut Muhammd Shubi Shalih di dalam buku Sejarah Al Qu'an karangan Athailah Al Qur'an (2010:14) adalah kalam yang mu'jiz (yang dapat melemahkan orang yang menentangnya) yang diturunkan kepada nabi Muhammad SAW yang tertulis dalam mushaf, yang disampaikan kita secara mutawatir yang membacanya dianggap ibadah.

Seseorang yang telah hafal Al Qur'an secara keseluruhan di luar kepala bisa disebut dengan juma' dan Huffazhul Qur'an. Pengumpulan Al Qur'an dengan cara menghafal (*Hifzuhu*) ini dilakukan pada masa awal penyiaran agama Islam, karena Al Qur'an pada waktu itu diturunkan melalui pendengaran. Pelestarian Al Qur'an melalui hafalan ini sangat tepat dan dapat dipertanggung jawabkan, mengingat Rasulullah SAW tergolong orang yang ummi (tidak lihai baca tulis) (1997:106). Dengan begitu dapat disimpulkan bahwa santri tahfidz adalah orang yang sedang mendalami atau sedang bersungguh-sungguh dalam menghafalkan Al Qur'an di suatu tempat lembaga pendidikan pondok pesantren.

Adapun tujuan akan didirikan kelas tahfidz diantaranya adalah sebagai berikut:

1. Menyelenggarakan kelas khusus Tahfidz Al Qur'an sebagai wadah untuk menumbuhkan generasi penghafal Al Qur'an.
2. Menyelenggarakan proses pendidikan, pengawasan, pembimbingan yang menjadikan Al Qur'an sebagai ruh sinergitas dengan keilmuan lainnya.
3. Mendoorng lahirnya generasi qur'ani yang memiliki aqidah yang benar, berprestasi, disiplin, mandiri dan berakhlak karimah (ta'zim dan tawadhu').
4. Program Akselerasi tahfidz
5. Menjadikan santri penghafal yang berkualitas
6. Membangun akhlak dengan Al Qur'an
7. Mencerdaskan generasi bangsa dengan Al Qur'an.

Santri kelas tahfidz dalam hal ini ialah semua santri yang mondok atau

menetap di lembaga ponpes tersebut yang mana santri-santri pilihan saja yang dipandang mampu dan mumpuni dalam mengikuti kelas tahfidz tingkat atas. Santri yang pilihan tersebut dengan memiliki kriteria mampu membaca dengan baik dan benar, mempunyai keinginan yang kuat untuk mengkaji dan menghafal Al Qur'an sesuai dengan target, memiliki kemampuan khusus dalam menghafal, dan siap untuk dibina selama pembelajaran. Adapun santri yang tidak berkesempatan untuk di kelas tahfidz tingkat atas maka akan difokuskan pada pelafalan dan perbaikan bacaan terlebih dahulu. Jika kelas Tahfidz tingkat atas 5:1 dalam menghafal, maka kelas regular 3:3, setengah untuk menghafal dan setengah lagi untuk perbaikan bacaan.

1.2.4 Ponpes Al Mizan Muhammadiyah Lamongan

A. Profil Ponpes Al Mizan Muhammadiyah Lamongan

Pondok pesantren Al Mizan asal mulanya adalah Panti Asuhan Muhammadiyah Cabang Lamongan yang didirikan pada tanggal 17 Agustus 1985 tepatnya di desa Banjarmendalan Kecamatan Lamongan Kabupaten Lamongan atau dikenal dengan alamat Jl Jenderal Sudirman No.1 (Utara Monumen Kadet Soewoko) Lamongan Jawa Timur. Panti Asuhan Muhammadiyah ini pada awalnya dirintis dan didirikan oleh Drs. HM Syukron (alm) yang kemudian didukung oleh Pimpinan Cabang Muhammadiyah Lamongan dan Aisyiyah Cabang Lamongan serta para tokoh / sesepuh Muhammadiyah Lamongan. Beliau [Drs. HM Syukron] adalah seorang mantan aktifis HM (Komisaris UII Surakarta), setelah tamat dari kuliahnya dia kembali ke Lamongan, Beliau memulai karir perjuangannya menjadi kepala PGAA Lamongan (1976-1979) yang sekarang menjadi Aliyah Pembangunan, Dosen/Dekan UNSURI Lamongan (1979-1982).

Setelah dibangunnya masjid At-Taqwa karena amanat beliau terpikirkan dengan ide-ide yang cemerlang diantaranya menginginkan adanya pengkaderan dengan system pondok pesantren dikalangan persyarikatan Muhammadiyah khususnya di Lamongan Kota, karena pada saat itu (sekitar tahun 80 an) di kota Lamongan untuk mencari seorang menjadi Ketua Cabang Muhammadiyah sangat sulit sekali. Dan di masjid At-Taqwa inilah beliau (HM Syukron) mendirikan

Madrasah Aliyah Muhammadiyah (1985) dan MTs Muhammadiyah (1986) yang diharapkan kelak menjadi tempat perkaderan bagi anak-anak khususnya dari kalangan Muhammadiyah dan umumnya umat Islam, baik sebagai kader ‘Ulama, kader Pemimpin ataupun Kader Mubaligh. Murid-murid tersebut merupakan cikal bakal dari anak asuh Panti Asuhan Muhammadiyah Cabang Lamongan, dimana murid-murid tersebut berasal dari keluarga kurnag mampu yang sebagian besar dititipkan di para aghniya’ yang lazim disebut Asuhan Keluarga.

Walaupun HM Syukron dari keluarga Nahdhiyin, namun setelah berguru di UII Surakarta dan HMI rupanya pola pikir dan langkah perjuangannya telah mengalami perubahan yang rasional dan modern. Seiring berjalannya waktu pembinaan yang berada dalam asuhan keluarga tengah mendapati berbagai kendala diantaranya tidak maksimalnya anak-anak dalam kegiatan belajar mengajar karena dengan alasan harus menyelesaikan tugas rumah dan lain-laain. Sehingga beliau bapak HM Syukron berangan-angan keras bagaimana cara mempunyai lahan yang nantinya dapat dibangun sebuah asrama. Alhamdulillah dengan bimbingan dan izin Allah SWT ada seorang Aghniya’ menyerahkan wakafnya dengan luas lahan 10 x 30 M kepada Muhammadiyah. Tanpa basa-basi beliau bapak HM Syukron mengusulkan supaya tanah tersebut tidak hanya dibangun musholla tetapi sekaligus asrama Panti Asuhan.



Gambar 2. 1 Gedung Ponpes Al Mizan

B. Program Tahfidzul Qur'an

Ponpes Al mizan yang kini semakin tampak di depan masyarakat Lamongan khususnya dan masyarakat Indonesia pada umumnya, memiliki berbagai program-program unggulan diantaranya Pekan Dakwah, Praktik Dakwah Lapangan, Paper, Praktek Studi Lapangan, I'tikaf, Kelas Bahasa dan Kelas Tahfidzul Qur'an. Berbagai program kegiatan yang dilaksanakan secara bertahap mulai dari tingkatan paling kecil hingga tingkatan yang tinggi.

Salah diatantara program-program tersebut ada beberapa program yang harus melalui tes terlebih dahulu yakni program kelas Tahfidz. Program kelas Tahfidz yang hanya menerima beberapa santri saja, karena keterbatasan sumber daya insani sehingga menjadikan kuota kelas tahfidz terbatas. Disisi lain juga karena faktor akan keefektifan dan keefesiensitas kegiatan akan lebih maksimal dalam pembelajarannya.

1.2.5 Bahasa Pemrograman PHP dan Notepad++

Bahasa pemrograman PHP merupakan bahasa pemrograman yang diciptakan oleh Dr. Leonardo Bernart pada tahun 1994. PHP singkatan dari Hypertext Preprocessor yang bersifat open source. Bahasa pemrograman PHP adalah bahasa pemrograman web yang bisa menyesuaikan tampilan konten tergantung situasi. Kelebihan bahasa pemrograman PHP yaitu pertama, merupakan bahasa script yang tidak perlu untuk dikompilasi (compile). Kedua, bahasa pemrograman juga bisa dijalankan diberbgaia system operasi.

PHP yakni dimana suatu bahasa pemrograman yang berbasis kode-kode yang digunakan sebagai pengolahan suatu data serta pengirimannya akan kembali ke web browser yang menjadi kode HTML. Selain itu PHP memberikan keuntungan seperti beberapa poin yang disuport oleh banyak komunitas. Hal seperti ini yang membuat PHP semakin berkembang. Selain itu juga, anda dapat belajar mengenai tips serta trik penggunaan dari berbagai komunitas secara lebih banyak lagi, lembaga pendidikan, maupun media sosial. Sebelum mempelajari PHP, adakalanya lakukan beberapa persiapan, seperti menyediakan web server dan program bantu editor PHP. (Oktavian, 2010)

Notepad ++ merupakan text editor suatu program, sehingga memungkinkan penggunaannya untuk bisa membuat tempalte web, design dan juga aplikasi tertentu. Dengan cara mengubah, membuat dan mengedit file text yang ada. Notepad ++ editor adalah text editor yang dipergunakan untuk Microsoft windos sehingga memungkinkan penggunanya dapat bekerja dengan banyak file terbuka dalam satu windows. Keunggulan Notepad++ jika dibandingkan dengan text editor lainnya diantaranya, pertama Notepad++ merupakan text editor yang sangat ringan. Kedua, text editor ini punya multi tab sehingga memungkinkan penggunanya membuka file secara bersamaan.

1.2.6 MySQL dan Xampp

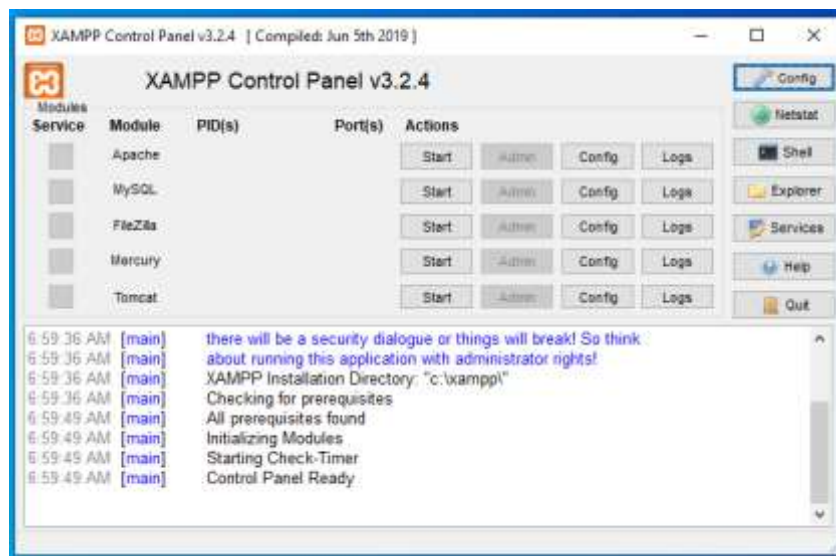
Dalam pembuatan program tentu dibutuhkan sebuah *database* untuk menyimpan data-data yang dibutuhkan dalam program tersebut. MySQL adalah sebuah program *database* yang bisa menerima dan mengirimkan data dengan cepat dengan menggunakan sintak SQL yang memudahkan seseorang dalam mengoperasikannya.

MySQL adalah database yang tersedia secara gratis bagi penggunaannya. MySQL gratis diakses oleh siapapun dalam keperluan pribadi maupun tidak tanpa harus membeli atau membayar lisensinya. MySQL pertama kali dirintis oleh *programmer* database bernama Michael Widenius. MySQL dapat diakses oleh seorang *client* maupun *server* sehingga MySQL dapat diartikan sebagai *database* yang bersifat *client-server*.

MySQL dan SQL tentu sangat berbeda arti. Arti dari MySQL sendiri adalah sebuah perangkat lunak *database*, sedangkan SQL adalah perintah-perintah atau sintak (*Query*) dalam MySQL. Adapun kelebihan dari MySQL ini adalah :

1. MySQL sebagai *Database Management System* (DBMS)
2. MySQL sebagai *Relation Database Management System* (RDBMS)
3. MySQL adalah *software database* yang gratis
4. MySQL merupakan *database* yang bersifat *client-server*
5. MySQL merupakan *database* dengan kapasitas penyimpanan yang sangat besar hingga mencapai ukuran *Gigabyte*.

Xampp adalah program aplikasi yang berfungsi sebagai server lokal untuk menampung berbagai jenis data website yang sedang dalam proses pengembangan (website online) sehingga tidak dapat diakses oleh banyak orang. Xampp sendiri diambil dari kata berikut. Pertama, X (Class Platform) merupakan aplikasi gratis yang bisa dikembangkan banyak orang. Ketiga, M (MySQL/MariaDB) merupakan database server yang dapat berfungsi untuk mengolah daftar melalui database. Keempat P (PHP) dimana menunjukkan eksistensinya menggunakan bahasa pemrograman php untuk membuat website dinamis. Dan yang terakhir adalah P (Perl) yang artinya bisa berjalan didalam banyak system operasi sehingga bersifat fleksibel.



Gambar 2.2 Koneksi Database dengan Xampp

1.2.7 PhpMyadmin

PhpMyadmin adalah sebuah aplikasi open source yang berfungsi untuk memudahkan manajemen MySQL. Dengan menggunakan PhpMyadmin kita dapat membuat database, membuat table, menginsert, menghapus dan mengupdate data dengan GUI dan terasa lebih mudah tanpa perlu mengetikkan perintah SQL secara manual.

1.2.8 Metode Naive Bayes

Algoritma Naive Bayes merupakan metode klasifikasi menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Algoritma Naive Bayes memperdiksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naive Bayes Classifier adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi/kejadian.

Naive Bayes Classifier bekerja sangat baik dibanding dengan model classifier lainnya. Hal ini dibuktikan pada jurnal Xhemali, Daniela, Chris J. Hinde, and Roger G. Stone. "Naive Bayes vs decision tree vs neural networks in the classification of training web pages."(2009), mengatakan bahwa "Naive Bayes Classifier memiliki tingkat akurasi yang lebih baik dibanding model classifier lainnya."

Keuntungan penggunaan adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (training data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Karena yang diasumsikan sebagai variabel independent, maka hanya variasi dari suatu variabel dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarians. Tahapan dari proses algoritma Naive Bayes adalah:

1. Menghitung jumlah kelas / label
2. Menghitung jumlah kasus per kelas
3. Kalikan semua variabel kelas
4. Bandingkan hasil per kelas

Kelebihan Naive Bayes adalah Mudah untuk dibuat dan Hasil yang bagus. Kekurangan Naive Bayes adalah asumsi independence antar atribut membuat akurasi berkurang (karena biasanya ada keterkaitan). Keuntungan dari menggunakan metode *Naive Bayes* adalah metode ini hanya menggunakan data pelatihan atau data *training*. Data *training* digunakan untuk menentukan parameter dalam sebuah proses pengklasifikasian. Naive Bayes bisa digunakan jauh lebih baik dalam situasi nyata dibanding dengan apa yang diharapkan.

Kekurangan Naïve Bayes sendiri diantaranya jika probabilitas kondisionalnya bernilai nol, maka probabilitas prediksi juga akan bernilai nol. Asumsi bahwa masing-masing variable independen membuat berkurangnya akurasi, karena biasanya ada korelasi anantara variable yang satu dengan variable yang lain. Yang mana keakuratannya tidak bisa diukur menggunakan satu probabilitas saja. Butuh bukti-bukti lain untuk membuktikannya. Sedang untuk membuat keputusan diperlukan pengetahuan awal atau pengetahuan mengenai masa sebelumnya. Keberhasilan sangat bergantung pada pengetahuan awal tersebut. banyak celah bisa mengurangi efektivitasnya. Dan untuk mendeteksinya berupa kata-kata saja, tidak bisa berupa gambar. (Bustami, 2013:127-146).

Persamaan dari metode *Naive Bayes* adalah sebagai berikut: (Pattekari:2012:290-294).

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \dots \dots \dots (2.1)$$

Dimana :

X : Data dengan class yang belum diketahui

C : Hipotesis data merupakan suatu class spesifik

P(C|X) : Probabilitas hipotesis C berdasar kondisi X (prosteriori probabilitas)

P(C) : Probabilitas hipotesis C

P(X|C) : Probabilitas X berdasarkan kondisi pada hipotesis

P(X) : Probabilitas X

Rumus diatas menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik karakteristik sampel pada kelas C (disebut juga likelihood), dibagi dengan peluang kemunculan karakteristik sampel secara global (disebut juga evidence). Karena itu, rumus diatas dapat pula ditulis sebagai berikut :

$$posterior = \frac{prior \times likelihood}{evidence} \dots \dots \dots (2.2)$$

Nilai Evidence selalu tetap untuk setiap kelas pada satu sampel. Nilai dari posterior tersebut nantinya akan dibandingkan dengan nilai nilai posterior kelas

lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus Bayes tersebut dilakukan dengan menjabarkan $(c|x_1, \dots, x_n)$ menggunakan aturan perkalian sebagai berikut :

$$P(C|X_1, \dots, X_n) = P(C)P(X_1, \dots, X_n|C) \dots\dots\dots (2.3)$$

$$= P(C)P(X_1|c)(X_2, \dots, X_n|C, X_1)$$

$$= P(C)P(X_1|c)P(X_2|C, X_1)(X_3, \dots, X_n|C, X_1, X_2)$$

$$= P(C)P(X_1|c)P(X_2|C, X_1)P(X_3|C, X_1, X_2) \dots P(X_n|C, X_1, X_2, \dots, X_{n-1})$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Disinilah digunakan asumsi independensi yang sangat tinggi (naif), bahwa masing masing petunjuk saling bebas (independen) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

$$= P(c|X_1, \dots, X_n) = P(C) \prod_{i=1}^n P(X_i|C) \dots\dots\dots (2.4)$$

$$= P(c|X) = P(X_1|c)P(X_2|c) \dots P(X_n|c)P(c) \dots\dots\dots (2.5)$$

Persamaan diatas merupakan model dari **Teorema Naive Bayes** yang selanjutnya akan digunakan dalam proses klasifikasi.

Untuk klasifikasi data kontinyu digunakan rumus Densitas Gauss :

$$P = (X_i = x_i | Y_i = y_i) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(x_i - u_{ij})^2}{2\sigma_{ij}^2}} \dots\dots\dots (2.6)$$

Keterangan :

1. P :Peluang
2. X_i :Atribut ke i
3. x_i :Nilai atribut ke i
4. Y :Kelas yang dicari
5. y_j :Sub kelas Y yang dicari
6. u :Mean, menyatakan rata rata dari seluruh atribut
7. σ : Deviasi standar, menyatakan varian dari seluruh atribut

Mean :

$$\mu = \frac{1}{n \sum_{i=1}^n x_i} \dots\dots\dots (2.7)$$

Deviasi Standar :

$$\sigma = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 \right]^{0.5} \dots\dots\dots (2.8)$$

Laplace Correction (Laplacian Estimator) atau **additive smoothing** adalah suatu cara untuk menangani nilai probabilitas 0 (nol). Dari sekian banyak data di training set, pada setiap perhitungan datanya ditambah 1 (satu) dan tidak akan membuat perbedaan yang berarti pada estimasi probabilitas sehingga bisa menghindari kasus nilai probabilitas 0 (nol). dimana nilai k adalah jumlah kelas atau bin dari atribut m_i .

$$P_i = \frac{m_i+1}{n+K} \dots\dots\dots (2.9)$$