

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Kajian Pustaka**

Kajian pustaka pada penelitian ini diambil dari sepuluh jurnal dengan rentang waktu 2017-2020. Penelitian untuk menghitung tingkat kecanduan sudah banyak dilakukan. Terdapat beberapa perbedaan pada setiap penelitian, seperti permasalahan yang ada, metode yang digunakan ataupun studi kasus yang ada pada penelitian. Berikut beberapa penelitian yang relevan dengan penelitian ini, antara lain:

Jurnal pertama yang dirujuk dalam penelitian ini diambil dari penelitian pada tahun 2017 yang dilakukan Yudi Prastyo, Puspa Eosina dan Fety Fatima. Variabel yang digunakan pada penelitian adalah *relapse*, *mood modification*, *conflict*, *tolerance* dan *cluster*/tingkat kecanduan. Hasil penelitian yang didapatkan yaitu terdapat empat kelompok kecanduan game *online* yaitu *cluster* dengan 12 data tingkat kecanduan berat, 21 data tingkat kecanduan sedang, 20 data tingkat kecanduan waspada dan 17 data tidak kecanduan game *online*. Kesimpulan yang diperoleh dari penelitian tersebut yaitu adanya korelasi antara tingkat kecanduan game *online* terhadap prestasi akademik mahasiswa di Universitas Ibn Khaldun Bogor [11].

Kemudian, jurnal kedua yang dirujuk dalam penelitian ini diambil dari penelitian pada tahun 2018 yang berjudul “*Fuzzy Logic* Menganalisis Pengaruh Media Sosial Terhadap Perilaku Masyarakat Adat Kenagarian Kinari”. Penelitian dengan menggunakan metode logika fuzzy Tahani ini bertujuan untuk mengetahui dampak media sosial terhadap perilaku masyarakat adat Kenagarian Kinari. Variabel pada penelitian ini yaitu variabel frekuensi dan variabel aktifitas. Selain variabel inputan terdapat variabel output kesimpulan. Dari hasil coba pada masyarakat dan tokoh masyarakat ditemukan tingkat akurasi yang cukup tinggi yaitu 85%. [12].

Jurnal ketiga yang dirujuk pada penelitian ini diambil dari penelitian yang berjudul “Pengaruh Media Sosial terhadap Prestasi Belajar Bahasa Inggris pada Program Studi Teknik Mesin ITN Malang”. Penelitian ini dilakukan oleh Tutut Nani Prihatmi pada tahun 2018. Penelitian ini bertujuan untuk mengetahui pengaruh dan uji korelasi dua variabel menggunakan metode regresi linear. Variabel pada penelitian ini yaitu durasi media sosial dan hasil belajar. Kesimpulan yang diperoleh dari penelitian yaitu penggunaan media sosial sangat berpengaruh terhadap perilaku belajar dan hasil belajar mahasiswa dikarenakan jika variabel durasi tinggi maka variabel hasil belajar rendah dan sebaliknya [13].

Jurnal keempat yang dirujuk pada penelitian ini diambil dari penelitian yang dilakukan oleh Alvendo Wahyu Aranski dan Sestri Novia Rizki pada 2018. Penelitian dengan logika fuzzy mamdani ini bertujuan untuk mendeteksi kecanduan penggunaan *smartphone* terhadap produktivitas kerja. Variabel pada penelitian ini yaitu variabel input spesifikasi *smartphone*, intensitas pengguna, statis sosial dan jenis pekerjaan serta output keputusan. Hasil penelitian yaitu logika fuzzy dapat menganalisa keputusan dalam memprediksi kecanduan penggunaan *smartphone* terhadap produktivitas kerja di kota Batam [14].

Selanjutnya, jurnal kelima yang dirujuk pada penelitian ini diambil dari penelitian yang dilakukan oleh Mochamad Farid Rifai, Hendra Jatnika dan Bowval Valentino pada tahun 2019. Penelitian ini berjudul “Penerapan Algoritma *Naive Bayes* Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi *Microsoft Office Specialist (MOS)*”. Penelitian ini digunakan untuk digunakan untuk klasifikasi kelulusan sertifikasi yaitu lulus dan tidak lulus. Variabel yang digunakan pada penelitian ini terdiri dari kriteria 1 hingga kriteria 5. Penelitian ini menghasilkan nilai probabilitas sebesar 0.001042 dengan tingkat akurasi yang tinggi yaitu 99% [15].

Kemudian, jurnal berikutnya yang dirujuk pada penelitian ini diambil dari penelitian yang dilakukan oleh M. Ibrahim, Efori Bu’ulolo, Ikwan Lubis

pada tahun 2020. Penelitian ini berjudul “Penerapan Algoritma *Naive Bayes Classifier* Untuk Mendeteksi Tingkat Kredibilitas *Hoax News/ Fake News* Pada Sosial Media Di Indonesia Berbasis Android (Studi Kasus: Kantor Tribun Medan)”. Penelitian dilakukan dengan menggunakan metode Algoritma *Naive Bayes Classifier*. Hasil penelitian ini adalah aplikasi sistem pendukung keputusan berbasis android yang dapat menentukan jenis berita bersifat *hoax* atau fakta [16].

Selanjutnya, terdapat jurnal penelitian yang berjudul “Implementasi Logika Fuzzy Mamdani Untuk Mengidentifikasi Tingkat Kecanduan Pelajar Terhadap *Game Online*”. Penelitian ini dilakukan oleh Nafasansono Harefa dan Murni Marbun pada tahun 2020. Metode yang digunakan pada penelitian ini adalah logika fuzzy Mamdani. Variabel yang digunakan pada penelitian ini adalah durasi waktu yang digunakan, *compulsion* (dorongan dari dalam diri sendiri) dan *withdrawal* (ketidaksanggupan menahan diri) serta variabel *output* yaitu tingkat kecanduan yang terdiri dari tidak candu, candu dan sangat candu. Hasil dari penelitian ini yaitu sistem yang dapat mengidentifikasi tingkat kecanduan pelajar terhadap *game online* [17].

Jurnal kedelapan yang dirujuk pada penelitian ini diambil dari penelitian yang berjudul “Klasifikasi Menggunakan Metode *Naive Bayes*: Tingkat Pengaruh Penggunaan Gadget terhadap Kematangan Kecerdasan Emosi Mahasiswa”. Penelitian ini dilakukan pada tahun 2020 oleh Umar Zaky, Ari Prasetyoaji dan Ilham Fathullah. Metode pada penelitian ini menggunakan algoritma *Naive Bayes Classifier*. Variabel pada penelitian antara lain durasi, konten dan tujuan. Hasil penelitian yaitu adanya klasifikasi tingkat pengaruh penggunaan *gadget* terhadap kematangan kecerdasan emosi mahasiswa. Hasil pengujian metode mendapatkan akurasi sebesar 82% [18].

Kemudian, terdapat jurnal penelitian yang berjudul “Sistem Pakar Klarifikasi Kecanduan Gadget Menggunakan Teori Arthurt T. Hovart dengan Metode *Naive Bayes Classifier* untuk Anak Sekolah Dasar”. Penelitian ini dilakukan oleh Luqman Affandi, Agung Nugroho Pramudhita dan Mardiana

Putri Sasmita pada tahun 2020. Penelitian dilakukan dengan menggunakan Algoritma *Naive Bayes Classifier*. Penelitian ini bertujuan untuk menentukan tingkat kecanduan *gadget* anak yaitu tipe kecanduan Rendah, sedang, dan berat beserta penanganan yang sesuai dengan tingkat kecanduannya. Berdasarkan 30 data yang ada, nilai akurasi untuk penelitian ini mempunyai hasil sebesar 86,67% [19].

Jurnal rujukan terakhir yang dirujuk dalam penelitian ini diambil dari penelitian yang meneliti tentang penerapan Algoritma *Naive Bayes* untuk mengklarifikasi pengaruh game *online* bagi remaja. Penelitian ini dilakukan oleh Nungky Asmiati dan Fatmawati pada tahun 2020. Data yang digunakan pada penelitian ini didapatkan melalui kuesioner *online* sebanyak 100 responden yang disebar melalui media sosial. Variabel yang digunakan dalam penelitian ini yaitu nama lengkap, jenis kelamin, usia, jumlah hari, durasi waktu, klarifikasi. Hasil perhitungan didapatkan kategori normal dengan akurasi sebesar 89.00% dan *Class recall* kecanduan 77.27%, *Class recall* normal 92.31%, *Class precision* kecanduan 73.91% dan *Class precision* normal 93.51% [20].

**Tabel 2.1 Perbandingan referensi**

No	Judul	Objek	Metode	Hasil	Perbedaan
1	Pembagian Tingkat Kecanduan <i>Game Online</i> Menggunakan <i>K-Means Clustering</i> serta Korelasinya terhadap Prestasi Akademik [11]	Data mahasiswa pada 6 Fakultas di Universitas Ibn Khaldun Bogor yaitu dari angkatan 2017 sampai dengan angkatan 2020	<i>K-Means Clustering</i>	Hasil penelitian yang didapatkan yaitu terdapat 4 kelompok kecanduan game <i>online</i> yaitu <i>cluster</i> yang tingkat kecanduan parah, tingkat kecanduan sedang, tingkat waspada dan tidak kecanduan. Nilai korelasi yang dihasilkan sebesar - 0,885 yang berarti bahwa adanya korelasi antara tingkat kecanduan game <i>online</i> terhadap prestasi akademik mahasiswa di Universitas Ibn Khaldun Bogor	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto menggunakan Algoritma <i>Naïve Bayes Classifier</i> . Penelitian menggunakan variabel <i>salience, mood modification, tolerance, withdrawal symptoms, conflict</i> , relapse, serta variabel <i>output</i> berupa tingkat kecanduan.

No	Judul	Objek	Metode	Hasil	Perbedaan
2	<i>Fuzzy Logic</i> Menganalisis Pengaruh Media Sosial Terhadap Perilaku Masyarakat Adat Kenagarian Kinari [12]	20 Orang Masyarakat Adat Kenagarian Kinari	Metode Fuzzy Tahani	Adanya sistem informasi <i>fuzzy logic</i> untuk menganalisa pengaruh media sosial terhadap perilaku masyarakat dengan tingkat keakuratan sebesar 85%	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto dengan menggunakan Algoritma <i>Naive Bayes Classifier</i>
3	Pengaruh Media Sosial terhadap Prestasi Belajar Bahasa Inggris pada Program Studi Teknik Mesin ITN Malang [13]	Angkatan 16 ITN Malang yaitu 70 mahasiswa Teknik Mesin	Metode regresi linier	Hasil penelitian menunjukkan bahwa penggunaan media sosial berpengaruh terhadap perilaku belajar dan hasil belajar mahasiswa di kampus	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto dengan menggunakan Algoritma <i>Naive Bayes Classifier</i>

No	Judul	Objek	Metode	Hasil	Perbedaan
4	Pemrograman Visual Basic Pendeteksi Dini Kecanduan Penggunaan <i>Smartphone</i> Terhadap Produktivitas Kerja Menggunakan Logika Fuzzy [14]	Masyarakat Kota Batam, Kepulauan Riau	Metode Fuzzy Mamdani	Hasil penelitian yaitu adanya sistem yang mampu membantu bagian kepegawaian untuk memprediksi tingkat kecanduan <i>smartphone</i> terhadap produktivitas kerja di kota Batam	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto dengan menggunakan Algoritma <i>Naive Bayes Classifier</i>
5	Penerapan Algoritma <i>Naive Bayes</i> Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi <i>Microsoft Office Specialist</i> (MOS) [15]	1024 data pada sertifikasi di laboratorium ITCC pada tahun 2017 – 2018	<i>Naive Bayes Classifier</i>	Hasil dari penelitian yaitu Algoritma <i>Naive Bayes Classifier</i> dapat memberikan informasi terkait prediksi peserta sertifikasi yang dinyatakan lulus dan tidak lulus. Penelitian ini memiliki tingkat akurasi sebesar 99%	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto. Penelitian menggunakan variabel <i>saliency, mood modification, tolerance, withdrawal symptoms, conflict, relapse,</i>

No	Judul	Objek	Metode	Hasil	Perbedaan
					serta variabel <i>output</i> berupa tingkat kecanduan.
6	Penerapan Algoritma <i>Naive Bayes Classifier</i> Untuk Mendeteksi Tingkat Kredibilitas <i>Hoax News/ Fake News</i> Pada Sosial Media Di Indonesia Berbasis <i>Android</i> (Studi Kasus: Kantor Tribun Medan) [16]	Data facebook dan twitter	Algoritma <i>Naive Bayes Classifier</i>	Hasil dari penelitian yaitu penentuan jenis berita bersifat <i>hoax</i> atau fakta dapat menggunakan metode <i>Naive Bayes Classifier</i> dapat digunakan dalam	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto dengan menggunakan Algoritma <i>Naive Bayes Classifier</i>
7	Implementasi Logika Fuzzy Mamdani Untuk Mengidentifikasi Tingkat Kecanduan Pelajar Terhadap Game <i>Online</i> [17]	Pelajar	Logika Fuzzy Mamdani	Hasil penelitian yaitu sistem yang menunjukkan apakah seorang pelajar termasuk dalam tingkatan tidak candu, candu atau sangat candu.	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto

No	Judul	Objek	Metode	Hasil	Perbedaan
8	Klasifikasi Menggunakan Metode <i>Naive Bayes</i> : Tingkat Pengaruh Penggunaan <i>Gadget</i> terhadap Kematangan Kecerdasan Emosi Mahasiswa [18]	Mahasiswa aktif pada jenjang pendidikan D1, D3, D4, S1, S2, S3, dan Profesi yang tersebar di provinsi D.I. Yogyakarta	<i>Naive Bayes</i>	Hasil pengujian metode mendapatkan akurasi sebesar 82%. Metode <i>naive bayes classifier</i> dapat mengklasifikasi tingkat pengaruh penggunaan <i>gadget</i> terhadap kematangan kecerdasan emosi.	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto dengan menggunakan Algoritma <i>Naive Bayes Classifier</i>
9	Sistem Pakar Klarifikasi Kecanduan <i>Gadget</i> Menggunakan Teori Arthurt T. Hovart dengan Metode <i>Naive Bayes Classifier</i> untuk Anak Sekolah Dasar [19]	30 data dari 2 Sekolah Dasar di kota Malang	<i>Naive Bayes Classifier</i>	Hasil dari penelitian ini yaitu dapat ditentukan tingkat kecanduan gadget Anak berdasarkan kecanduan rendah, sedang, dan berat menggunakan <i>naive bayes classifier</i> . Penelitian ini mempunyai nilai akurasi sebesar 86,67%	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto dengan Algoritma <i>Naive Bayes Classifier</i> . Penelitian menggunakan variabel salience, mood

No	Judul	Objek	Metode	Hasil	Perbedaan
					modification, tolerance, withdrawal symptoms, conflict, relapse, serta variabel output berupa tingkat kecanduan
10	Penerapan Algoritma <i>Naive Bayes</i> Untuk Mengklasifikasi Pengaruh Negatif <i>Game Online</i> Bagi Remaja Milenial [20]	100 data pengguna <i>game online</i>	<i>Naive Bayes</i>	Hasil perhitungan <i>Naive Bayes</i> didapatkan kategori normal dengan akurasi sebesar 89.00% dan <i>Class recall</i> kecanduan 77.27%, <i>Class recall</i> normal 92.31%, <i>Class precision</i> kecanduan 73.91% dan <i>Class precision</i> normal 93.51%	Penelitian yang akan dilakukan untuk menganalisis tingkat kecanduan media sosial instagram dan tiktok mahasiswa IT Telkom Purwokerto dengan menggunakan Algoritma <i>Naive Bayes Classifier</i>

Berdasarkan beberapa penelitian yang ada, peneliti memutuskan untuk melakukan penelitian tentang analisis tingkat kecanduan media sosial instagram dan tiktok menggunakan Algoritma *Naive Bayes Classifier*. Peneliti menggunakan Algoritma *Naive Bayes Classifier* karena algoritma *Naive Bayes Classifier* cepat dalam proses perhitungan, memiliki algoritma yang sederhana dan akurasi yang tinggi.

## **2.2. Landasan Teori**

### **2.2.1. Media Sosial**

Media sosial adalah media *online* yang dimana pengguna dapat dengan mudah berpartisipasi, berbagi, dan menciptakan isi meliputi blog, jejaring sosial, wiki, forum dan dunia virtual. Media sosial juga dapat dikatakan sebagai media *online* yang mendukung interaksi sosial dengan menggunakan teknologi yang mengubah komunikasi menjadi dialog interaktif [21].

Media sosial yang ada saat ini antara lain Facebook, WhatsApp, Twitter, Instagram, Youtube, Tiktok dan sebagainya. Media sosial yang sedang ramai digunakan saat ini adalah instagram dan tiktok Media sosial yang sedang ramai digunakan saat ini adalah instagram dan tiktok. Berdasarkan riset *AppTopia*, TikTok dan instagram menjadi media sosial yang paling banyak diunduh pada tahun 2021. Tiktok menempati posisi pertama dengan 656 juta unduhan sedangkan instagram menempati posisi kedua dengan 545 juta unduhan [3].

Jika media tradisional menggunakan media cetak dan media *broadcast*, maka media sosial menggunakan internet. Media sosial digunakan sebagai sarana komunikasi untuk membuat para penggunanya berpartisipasi secara aktif dengan memberi kontribusi dan *feedback* secara terbuka, baik untuk membagi informasi maupun memberi respon secara *online* dalam waktu yang cepat.

Media sosial memiliki pengaruh positif dan negatif bagi penggunanya. Beberapa pengaruh media sosial antara lain:

A. Dampak positif media sosial:

1. Dapat mempermudah komunikasi jarak jauh dengan orang lain.
2. Menambah wawasan berpikir dan pengetahuan melalui media sosial
3. Memfasilitasi pembelajaran siswa, seperti menemukan materi, mempelajari hal-hal baru, dan lain-lain.
4. Dapat memberikan informasi yang akurat dan tepat, seperti informasi tentang perguruan tinggi, lowongan pekerjaan atau beasiswa dan lain-lain.

B. Dampak negatif media sosial:

1. Menimbulkan sifat candu, lupa waktu, dan mengabaikan pekerjaan.
2. Maraknya pornografi di media sosial/internet,
3. Banyaknya berita bohong/*hoax*
4. Membuat penggunaannya menjadi malas [22].

### 2.2.2. Kecanduan media sosial

Kecanduan media sosial adalah gangguan psikologis dimana penggunaannya menambah jumlah penggunaan sehingga membangkitkan kesenangan yang dapat menyebabkan gangguan kecemasan dan perubahan *mood*, gangguan afeksi (depresi, kesulitan dalam menyesuaikan diri) dan terganggunya kehidupan sosial [5]. Pengguna yang kecanduan media sosial pastilah menunjukkan kriteria-kriteria tertentu.

Kecanduan media sosial memiliki enam aspek yang saling berhubungan. Seseorang dikatakan kecanduan media sosial apabila memenuhi tiga dari aspek tersebut. Aspek-aspek kecanduan media sosial, antara lain:

1. *Salience*

Hal ini terjadi ketika media sosial mendominasi pikiran pengguna. Pengguna akan selalu memikirkan media sosial, bahkan ketika mereka tidak sedang mengaksesnya.

2. *Mood modification*

Individu yang mendapat kepuasan saat menggunakan media sosial, perubahan mood.

3. *Tolerance*

Terjadinya peningkatan jumlah penggunaan untuk mencapai kepuasan sehingga aktifitas penggunaan yang melebihi waktu.

4. *Withdrawal*

Perasaan yang terjadi ketika penggunaan media sosial dikurangi, tidak tenang jika belum membuka media sosial

5. *Conflict*

Konflik antara pengguna dengan lingkungan sekitar, dengan tugas atau pekerjaan, dengan diri sendiri dikarenakan terlalu banyak menggunakan media sosial

6. *Relapse*

Individu yang telah mencoba tetapi gagal untuk mengendalikan, mengurangi penggunaan media sosial [6].

### 2.2.3. Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi yang pengetahuan yang terkait dari berbagai *database* besar [23]. Data mining juga disebut sebagai *Knowledge in Discovery Database* (KDD) telah digunakan diberbagai bidang seperti bisnis, e-commerce, astronomi, geografi, kesehatan dan pendidikan.

Data mining dibagi menjadi beberapa kelompok, antara lain:

1. Prediksi

Memprediksi hasil yang akan ada dimasa mendatang

Contoh: Prediksi kenaikan harga sembako 2 bulan yang akan datang

2. Estimasi

Mengestimasi dengan variabel target numerik

3. Klasifikasi

Menemukan model yang menjelaskan atau membedakan kelas data untuk memperkirakan kelas suatu objek yang belum diketahui

Contoh: Menentukan seseorang lulus atau tidak

4. Pengelompokan

Menghasilkan label kelas untuk sebuah kelompok data.

5. Asosiasi

Menghitung hubungan antar dua atau lebih atribut [24].

#### 2.2.4. *Naive Bayes Classifier*

*Naive Bayes Classifier* adalah klarifikasi yang berdasarkan pada teorema Bayes. *Naive Bayes Classifier* digunakan untuk menghitung probabilitas setiap kelas dengan asumsi bahwa tidak ada ketergantungan antara satu kelas dengan kelas lainnya [23]. *Naive Bayes Classifier* mempunyai kecepatan dan yang tinggi jika digunakan pada *database* yang besar [25].

Tahapan dari algoritma *Naive Bayes Classifier* adalah:

1. Menghitung jumlah kelas
2. Menghitung jumlah kasus per kelas
3. Mengalikan semua variabel kelas
4. Membandingkan hasil berdasarkan kelas

Rumus *Naive Bayes Classifier* dapat dilihat pada persamaan 2.1 sebagai berikut:

$$P(H | X) = \frac{P(X | H).P(H)}{P(H)} \quad (2.1)$$

Keterangan:

X : data dengan *class* yang belum diketahui

H : hipotesis data menggunakan suatu *class* spesifik

P(H|X) : probabilitas hipotesis H berdasar kondisi X

P(H) : Probabilitas hipotesis H (prior probabilitas)

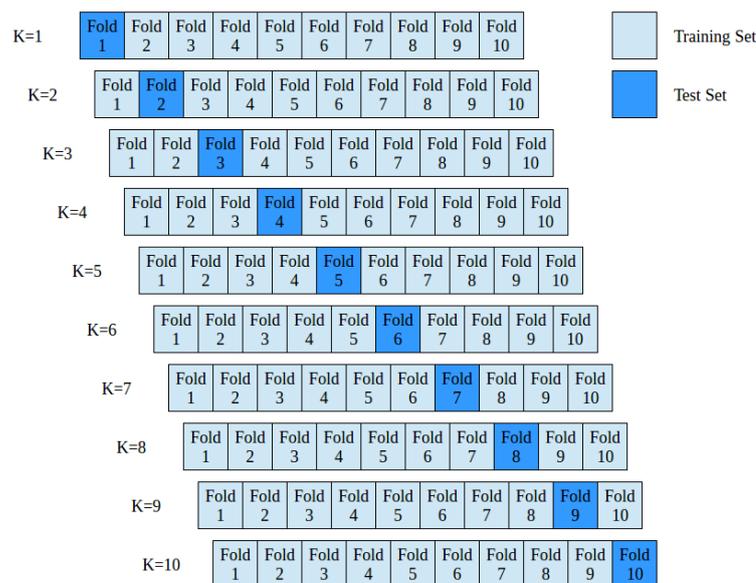
$P(X|H)$  : Probabilitas X berdasar kondisi pada hipotesis H

$P(X)$  : Probabilitas H

### 2.2.5. Cross Validation

*Cross validation* adalah metode statistik yang dapat digunakan untuk mengevaluasi kinerja model atau algoritma yang dimana data dipisahkan menjadi dua subset, yaitu data proses pembelajaran dan data validasi/evaluasi [26]. Pada *cross validation*, setiap *record* digunakan beberapa kali dalam jumlah yang sama untuk data latih dan untuk data uji.

*K-fold cross validation* adalah metode statistik yang digunakan untuk evaluasi kinerja yang lebih stabil dan menyeluruh daripada menggunakan pemisahan dataset menjadi satu set pelatihan dan pengujian. Dalam *K-fold cross validation*, data akan dibagi beberapa bagian berulang kali (iterasi) dan kemudian model dilatih. Nilai dari parameter K menentukan seberapa banyak *dataset* yang akan dibagi dan seberapa banyak iterasi *training* dan *testing* yang akan diproses [27]. Adapun diagram k-fold cross validation dapat dilihat pada Gambar 2.1.



**Gambar 2.1. Diagram K-fold Cross Validation**

### 2.2.6. Confusion Matrix

*Confusion Matrix* adalah alat yang berguna untuk menganalisis seberapa baik *classifier* mengenali *tuple* dari kelas yang berbeda. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar [28]. Confusion matrix digambarkan dengan tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan [29]. Confusion matrix digunakan untuk mengetahui nilai dari accuracy, precision, recall, dan f1-score [30]. Tabel *confusion matrix* dapat dilihat pada Tabel 2.2.

**Tabel 2.2. Multi Class Confusion Matrix**

		Prediksi		
		Positif	Negatif	Netral
Aktual	Positif	TPos	FposNeg	FposNet
	Negatif	FNegPos	Tneg	FNegNet
	Netral	FNetpos	FnetNeg	TNet

Perhitungan Akurasi dapat dilihat pada persamaan 2.3 sebagai berikut:

$$Akurasi = \frac{TPos + TNeg + TNet}{TPos + FPosNeg + FPosNet + FNegPos + TNeg + FNegNet + FNetPos + FNetNeg + FNet} * 100\% \quad (2.3)$$

Keterangan:

- TPos : jumlah prediksi benar pada data dengan label benar.
- TNeg : jumlah prediksi benar pada data dengan label salah.
- TNet : jumlah prediksi benar pada data dengan label Netral.
- FPosNeg : jumlah prediksi salah pada data dengan label benar dan salah.
- FPosNet : jumlah prediksi salah pada data dengan label benar dan netral.
- FNegPos : jumlah prediksi salah pada data dengan label salah dan benar.

- FNegNet : jumlah prediksi salah pada data dengan label salah dan netral.
- FNetPos : jumlah prediksi salah pada data dengan label netral dan benar.
- FNetNeg : jumlah prediksi salah pada data dengan label netral dan salah.