

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*Data mining* atau disebut juga dengan *knowledge discovery in database* (KDD) adalah proses mengolah data yang berukuran besar dengan mengekstrak informasi yang diperlukan dan menemukan pengetahuan tersembunyi dalam suatu database sehingga mendapatkan nilai tambah yang berguna dalam pengambilan keputusan. (Wu dkk, 2021; Phazaniraja & Sountharajan, 2020; Kaushik dkk, 2021; Tan & Lin, 2021; Xiao & Hu, 2021; Fote dkk, 2020). *Data mining* adalah proses penggalian informasi dan pengetahuan tentang nilai potensial dari data yang besar, acak, kabur, tidak lengkap, dan mengandung *noise* (Yu, 2021). Proses *data mining* ini meliputi pembersihan data, integrasi data, pemilihan data, transformasi data, penambahan data, evaluasi pola, dan representasi pengetahuan (Phazaniraja & Sountharajan, 2020). *Data mining* berperan penting dalam membantu banyak ilmuwan dan praktisi memahami bagaimana mengelola data untuk mengekstrak informasi yang berguna. (Martarelli & Nagano, 2022). Pengetahuan yang dihasilkan dari proses *data mining* dapat berkontribusi untuk perbaikan berbagai bidang (Vucetic dkk, 2020).

*Data mining* bertujuan untuk menemukan informasi yang berguna pada database untuk lebih memahami masa lalu, atau bahkan memprediksi masa depan (Fatemi dkk, 2020). *Data mining* merupakan disiplin penting yang dapat memberikan keuntungan signifikan di bidang penemuan pengetahuan dan analitik bisnis untuk institusi modern (Vucetic dkk, 2020). Asosiasi, *clustering*, klasifikasi, dan prediksi adalah beberapa teknik data mining yang sering digunakan (Kaushik dkk, 2021; Rao dkk, 2021; Singh dkk, 2021).

*Association Rule Mining* (ARM) adalah skema paling umum yang digunakan untuk mengenali pola tersembunyi dari data transaksional yang berukuran besar. ARM

mengenali pola, korelasi, struktur dan asosiasi dari kelompok item. Pendekatan ARM yang tersedia ada berbagai macam, diantaranya adalah penghitungan itemset dinamis, apriori, pencarian *pincher*, partisi dan algoritma *frequent pattern tree* (Sharmila & Vijayarani, 2020). ARM banyak digunakan dalam berbagai bidang, diantaranya *market basket analysis*, diagnosis medis, bioinformatika, layanan IoT, analisis perilaku user, aplikasi smartphone, dan keamanan siber (Kaushik dkk, 2021; Domadiya & Rao, 2021; Sarker, 2021).

Algoritma apriori adalah algoritma yang pertama kali diusulkan untuk menemukan aturan asosiasi, yang menggunakan model *generate-and-test* untuk mengidentifikasi kandidat yang mungkin pada setiap tingkatan. Algoritma apriori terdiri dari dua tahap operasi, yaitu menemukan *frequent items*, dan kemudian menambang aturan asosiasi. (Wu dkk, 2021; Liao dkk, 2020). *Frequent items* adalah kumpulan item yang sering muncul bersama, diukur dengan parameter *Support*. Aturan asosiasi diukur dengan dua parameter, yaitu *Support* dan *Confidence*. *Support* adalah batas minimum ditentukan untuk pemilihan *frequent itemsets* dan aturan asosiasi. *Confidence* adalah keakuratan aturan asosiasi yang digunakan untuk menentukan peringkat aturan (Zeng & Xiao, 2020; Yazdani dkk, 2021).

Efisiensi algoritma apriori dasar yang belum dimodifikasi tergolong rendah untuk dataset yang berukuran besar karena membutuhkan banyak *resource* dalam melakukan *scanning* basis data secara berulang untuk menemukan *frequent itemset*. Metode alternatif yang lebih efisien dibanding algoritma apriori dasar telah ditemukan, antara lain Apriori-TID, FP-Growth, ECLAT, H-Mine, dll. (Borah & Nath, 2020; Li & Sheu, 2021; Ahmed, 2021; Bashir, 2019). Algoritma apriori dapat ditingkatkan efisiensinya jadi lebih baik dengan menggunakan metode MapReduce saat melakukan *scanning* database yang berukuran besar untuk menemukan *frequent itemsets*. MapReduce memungkinkan proses pencarian *frequent itemsets* dapat dilakukan secara parsial dan paralel sehingga dapat mempersingkat waktu pemrosesan (Wang & Gao, 2021; Raj dkk, 2020; Lou dkk, 2020). Penerapan *data mining* di dunia nyata, terutama algoritma apriori, dapat ditemukan di berbagai bidang, diantaranya pemasaran, diagnostik medis, penambangan web, analisis lalu lintas, teknologi informasi dan forensik. (Datta dkk, 2020).

Hasil penelitian yang telah dilakukan peneliti sebelumnya pada data penderita kanker paru-paru, algoritma apriori dapat mengidentifikasi aturan potensial antara kanker paru-paru dan gejalanya. Algoritma apriori diterapkan untuk menganalisis dan

mempelajari faktor-faktor patogen kanker paru-paru berdasarkan data rekam medis elektronik kanker paru-paru dari departemen onkologi rumah sakit selama bertahun-tahun. Algoritma apriori dapat merumuskan dengan baik hubungan antara kanker paru-paru dan faktor-faktor patogen, yang dapat membantu dokter membuat diagnosis dengan cepat, akurat dan efisien, memiliki signifikansi klinis yang penting untuk pencegahan awal dan pengobatan kanker paru-paru (Guo dkk, 2021).

Hasil penelitian lain pada serangan jantung, algoritma apriori dapat dimanfaatkan untuk memprediksi serangan jantung yang merupakan penyebab kematian utama di banyak negara. Pasien yang menderita penyakit jantung memerlukan diagnosis cepat, pengobatan dini dan pengamatan terus-menerus. bahwa Algoritma apriori yang digunakan berhasil menemukan aturan asosiasi prediksi penyakit jantung yang dikonfirmasi sebagai aturan yang valid oleh ahli penyakit jantung, dengan skor kepercayaan tertinggi 98% dalam memprediksi penyakit jantung (Yazdani dkk, 2021).

Hasil penelitian pada Tambang Pingdingshan di Provinsi Henan, China algoritma apriori juga telah digunakan untuk meminimalisir dampak bencana. Studi kasus pada data bencana ledakan batubara dan gas di Tambang Pingdingshan menunjukkan bahwa algoritma apriori dapat menemukan aturan asosiasi penyebab ledakan batubara dan gas. Informasi tersebut sangat berguna dalam pencegahan dan pengelolaan ledakan batubara dan gas (Xie dkk, 2019).

Hasil penelitian di bidang transportasi, algoritma apriori dapat mengidentifikasi penyebab utama kecelakaan jalan tol yang melibatkan kendaraan pengangkut material berbahaya, menarik hubungan antara kecelakaan dan faktor risikonya, sehingga memberikan hasil yang lebih mudah dipahami untuk peningkatan keselamatan di jalan tol. Kecelakaan yang melibatkan kendaraan pengangkut *Hazardous Material* (HAZMAT) dapat menimbulkan korban jiwa dan kerugian materi yang besar sehingga perlu dilakukan identifikasi terhadap penyebab utama kecelakaan dan membuat tindakan pencegahan. Studi kasus dilakukan menggunakan data kecelakaan yang diperoleh dari database Korea Expressway Corporation dari 2008 hingga 2017. ARM dilakukan menggunakan algoritma Apriori, dan total 855 aturan menarik dihasilkan. Hasil penelitian menunjukkan bahwa kecelakaan yang melibatkan kendaraan pengangkut HAZMAT sangat terkait dengan pengemudi pria, kecelakaan yang melibatkan kendaraan tunggal, kondisi cuaca cerah, siang hari, dan segmen jalur utama (Hong dkk 2020).

Hasil penelitian yang membandingkan algoritma apriori dengan *collaborative filtering* untuk menganalisis preferensi pengguna pada penyewaan film MovieLens, menunjukkan bukti bahwa algoritma apriori mengungguli algoritma *collaborative filtering* pada basis data MovieLens berdasarkan akurasi prediksi peringkat (Singh dkk, 2021).

Hasil penelitian pada bidang akademik menunjukkan algoritma Apriori untuk mengidentifikasi aturan asosiasi yang dapat digunakan pada penilaian bakat untuk promosi staf akademik di universitas lokal. Data yang diolah adalah database akademik perguruan tinggi. Penelitian berhasil menemukan aturan asosiasi yang relevan dan signifikan dalam penilaian bakat staff (Othman dkk, 2019).

Hasil penelitian di Dubai, Uni Emirat Arab, algoritma apriori dapat digunakan untuk menganalisis data kecelakaan lalu lintas pada tahun 2017. Aturan asosiasi yang ditemukan menunjukkan bahwa penyebab kecelakaan terbesar pada akhir pekan adalah pengemudi yang dalam keadaan mabuk, sedangkan pada hari-hari biasa adalah kurangnya jaga jarak antar kendaraan (John & Shaiba, 2019).

Perawat adalah tenaga profesional pemberi asuhan di bidang kesehatan. Menurut Undang-Undang No 38 Tahun 2014 perawat adalah seseorang yang telah lulus pendidikan tinggi baik di dalam maupun luar negeri diakui pemerintah sesuai dengan ketentuannya. Kompetensi (*Competence*) adalah gambaran hal yang harus dilakukan seseorang agar dapat melaksanakan pekerjaannya dengan baik (Pebriani dkk, 2021). Perawat adalah sebuah profesi yang berbasis pada kompetensi yang melaksanakan asuhan keperawatan. Seorang perawat dituntut untuk memenuhi kualifikasinya dengan memiliki Surat Tanda Registrasi (STR) sebagai bentuk pengakuan Majelis Tenaga Kesehatan Indonesia (MTKI). STR tersebut diperoleh dengan syarat memiliki sertifikat kompetensi yang didapatkan setelah lulus mengikuti uji kompetensi. (Hadi dkk, 2020)

Uji kompetensi adalah proses pengukuran pengetahuan, keterampilan, dan perilaku peserta didik pada perguruan tinggi yang menyelenggarakan program studi keperawatan. Tujuan dilakukannya uji kompetensi khususnya terhadap perawat lulusan baru adalah untuk melindungi masyarakat dengan memberikan jaminan bahwa perawat pada *entry level registered* memiliki kompetensi yang dipersyaratkan untuk dapat menjalankan praktek profesi secara aman dan efektif (Palingrungi dkk, 2021).

Tingkat kelulusan mahasiswa keperawatan dalam mengikuti Uji Kompetensi Ners Indonesia (UKNI) sejak periode ke-IV terus menerus menurun dan sangat rendah

jika dibandingkan dengan beberapa negara yang tingkat kelulusannya diatas 80% (Palingrungi dkk, 2021; Choeron & Metrikayanto, 2020). Hasil evaluasi yang menunjukkan bahwa dari tahun ke tahun tingkat kelulusan Uji Kompetensi Perawat tidak mengalami kenaikan yang signifikan (Pebriani dkk, 2021).

Peserta UKNI yang tidak lulus perlu mendapatkan perhatian, karena apabila tidak direspon dengan cepat dan tepat maka akan berdampak terhadap kualitas lulusan yang dinilai tidak kompeten sebagai Ners, mutu dari perguruan tinggi akan diragukan oleh masyarakat serta banyaknya dampak negatif terhadap mahasiswa Ners yang tidak lulus uji kompetensi. Ketidaklulusan mahasiswa keperawatan pada UKNI juga dapat menyebabkan mahasiswa tidak bisa memiliki Surat Tanda Registrasi (STR), sehingga bisa menyulitkan lulusan untuk mendapatkan pekerjaan di pelayanan kesehatan dan melalui STR tersebut perawat dapat melakukan praktik mandiri sebagai tanda pengakuan terhadap kompetensi perawat yang telah lulus Uji Kompetensi (Palingrungi dkk, 2021).

Kelulusan uji kompetensi perawat dapat dipengaruhi dari berbagai faktor. Faktor yang memiliki pengaruh signifikan terhadap kemampuan peserta untuk dapat lulus uji kompetensi adalah persiapan yang matang, kesempatan untuk menjelaskan pemaparan terhadap praktik klinik, lingkungan belajar dan metode yang digunakan dalam pembelajaran klinik (Pebriani dkk, 2021). Prediktor yang mempengaruhi keberhasilan seseorang dalam menghadapi ujian kompetensi yaitu prestasi akademik, hasil *try out*, bimbingan intensif, status akreditasi, *independent test taker* berbasis web *e-learning* dan faktor demografi. Hubungan antar prediktor kelulusan yang diidentifikasi tersebut dapat diteliti lebih lanjut dengan teknologi data mining menggunakan algoritma apriori (Palingrungi dkk, 2021).

Penelitian ini mengidentifikasi hubungan antar prediktor kelulusan uji kompetensi perawat serta menghitung tingkat akurasi hubungan tersebut. Prediktor yang digunakan adalah IPK, akreditasi perguruan tinggi, hasil *try out*, tingkat partisipasi dalam bimbingan intensif, dan penguasaan aspek-aspek tertentu pada materi keperawatan. Tujuan penelitian adalah mengetahui faktor-faktor yang harus diperhatikan oleh lulusan atau perguruan tinggi keperawatan untuk persiapan yang lebih matang dalam menghadapi uji kompetensi. Berdasarkan uraian tersebut penulis melakukan penelitian dalam bentuk tesis dengan judul **Akurasi dalam Analisis Kompetensi Calon Tenaga Keperawatan Menggunakan Algoritma Apriori**.

## **1.2 Perumusan Masalah**

Berdasarkan permasalahan yang ada, supaya tesis ini sesuai dengan tujuan yang ingin dicapai, maka penulis merumuskan beberapa permasalahan sebagai berikut:

1. Bagaimana penerapan algoritma apriori dapat menemukan faktor prediktor yang berpengaruh terhadap kompetensi calon tenaga keperawatan?
2. Bagaimana penerapan algoritma apriori dapat menemukan aturan asosiasi antar faktor prediktor kompetensi perawat?
3. Bagaimana akurasi aturan asosiasi yang ditemukan oleh algoritma apriori?

## **1.3 Batasan Masalah**

Berdasarkan latar belakang masalah yang telah dipaparkan diatas dan pembahasan pada penelitian ini tidak menyimpang maka penulis membatasi ruang lingkup objek penelitian. Penelitian ini hanya berfokus pada:

1. Ruang lingkup penelitian ini hanya menganalisis data peserta bimbingan belajar APPSKEP.
2. Algoritma apriori yang diterapkan hanya untuk menganalisis faktor prediktor kelulusan yang ada pada peserta bimbingan belajar APPSKEP yang telah mengikuti UKNI.
3. Data yang akan diolah dalam bentuk basis data MySQL.
4. Data diolah di sistem yang dibangun khusus untuk penelitian ini menggunakan PHP.

## **1.4 Tujuan Penelitian**

Berdasarkan latar belakang dan hasil dari penelitian sebelumnya maka dapat dirumuskan beberapa tujuan yang akan diraih dalam melakukan penelitian ini, adapun tujuan yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma apriori untuk menemukan faktor prediktor yang berpengaruh terhadap kompetensi calon tenaga keperawatan
2. Menerapkan algoritma apriori untuk menemukan aturan asosiasi antar faktor prediktor kompetensi calon tenaga keperawatan.
3. Menguji akurasi aturan asosiasi yang dihasilkan oleh algoritma apriori.

## **1.5 Manfaat Penelitian**

Penelitian ini bertujuan untuk dapat memberikan manfaat yang dapat berguna bagi berbagai pihak dan penelitian selanjutnya. Adapun manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Memberikan pengetahuan relevan kepada calon tenaga keperawatan tentang faktor prediktor yang terkait dengan kondisi masing-masing individu dalam kompetensi sesuai UKNI sehingga calon peserta dapat mempersiapkan diri lebih matang dalam menghadapi UKNI.
2. Memberikan pengetahuan yang relevan kepada APPSKEP tentang faktor prediktor terkait dengan demografi user dalam kompetensi tenaga keperawatan agar dapat mengoptimalkan sistem pembelajaran.
3. Memberikan pengetahuan yang relevan kepada perguruan tinggi tentang faktor prediktor terkait dengan kampus dan mahasiswanya dalam uji kompetensi tenaga keperawatan agar dapat menjaga dan meningkatkan kualitas lulusan perguruan tinggi.

## **1.6 Sistematika Penulisan**

Sistematika yang akan menjadi panduan dalam penulisan laporan penelitian Data Mining dengan algoritma apriori dalam analisis kompetensi calon tenaga keperawatan adalah sebagai berikut:

**Bab I PENDAHULUAN**

Bab ini berisikan tentang latar belakang penelitian, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penelitian.

**Bab II LANDASAN TEORI**

Bab ini berisikan penerapan algoritma apriori serta teori-teori tentang Data Mining yang berhubungan dalam penelitian yang dilakukan.

**Bab III METODE PENELITIAN**

Bab ini berisikan kerangka kerja, mendefinisikan ruang lingkup, analisa masalah, menentukan tujuan, mempelajari literatur, mengumpulkan data dan informasi, menganalisa dan menerapkan algoritma apriori, mengimplementasikan algoritma apriori, menguji data dan menarik kesimpulan.

**Bab IV ANALISA DAN PERANCANGAN SISTEM**

Bab ini berisikan analisa dari penelitian yang dilakukan, analisa kebutuhan sistem dan analisa dari perancangan sistem yang dibuat dan penerapan algoritma apriori pada Data Mining dalam analisis kompetensi calon tenaga keperawatan.

**Bab V IMPLEMENTASI SISTEM**

Bab ini berisikan proses implementasi dari hasil analisa dan perancangan sistem Data Mining menggunakan algoritma apriori yang telah dibuat dan hasil yang telah berjalan.

**Bab VI KESIMPULAN DAN SARAN**

Bab ini berisikan kesimpulan dari bab-bab yang sudah dicantumkan mengenai hasil penerapan algoritma apriori. Bab ini juga berisi saran yang diharapkan yang dapat digunakan untuk penelitian selanjutnya.