

**PENENTUAN POLA PELANGGARAN SANTRI PADA BAGIAN
UBUDIYAH MENGGUNAKAN ALGORITMA APRIORI
(Studi Kasus: Pondok Pesantren Salafiyah Syafi'iyah Sukorejo)**

SKRIPSI



Oleh :

SYAMSUDDIN

2020502030

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2024

**PENENTUAN POLA PELANGGARAN SANTRI PADA BAGIAN
UBUDIYAH MENGGUNAKAN ALGORITMA APRIORI
(Studi Kasus: Pondok Pesantren Salafiyah Syafi'iyah Sukorejo)**

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Persyaratan dalam Menyelesaikan Program
Sarjana (S-1) pada Program Studi Sistem Informasi
Fakultas Sains dan Teknologi Universitas Ibrahimi

Oleh :

SYAMSUDDIN

2020502030

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMI
SITUBONDO**

2024

HALAMAN PERSETUJUAN

Nama : Syamsuddin
NPM : 2020502030
Judul : **Penentuan Pola Pelanggaran Santri Pada Bagian Ubudiyah
Menggunakan Algoritma Apriori (Studi Kasus: Pondok
Pesantren Salafiyah Syafi'iyah Sukorejo)**

Oleh :

SYAMSUDDIN

2020502030

Telah disetujui oleh:

Pembimbing 1

Muhammad Ali Ridla, M.Kom.
NIDN: 728089402

Pembimbing 2

Akhlis Munazilin, S.Kom., M.T
NIDN: 715078902

HALAMAN PENGESAHAN
SKRIPSI

PENENTUAN POLA PELANGGARAN SANTRI PADA BAGIAN
UBUDIYAH MENGGUNAKAN ALGORITMA APRIORI
(Studi Kasus: Pondok Pesantren Salafiyah Syafi'iyah Sukorejo)

SYAMSUDDIN

2020502030

Telah dipertahankan di depan dewan penguji Sidang/Munaqosyah Skripsi pada hari Senin , 12 Agustus 2024 sebagai salah satu syarat memperoleh gelar Sarjana (S.Kom) pada Fakultas Sains dan Teknologi Universitas Ibrahimiy.


Tim Penguji,

Ketua Sidang,



Ahmad Homaidi, M.Kom.
NIDN: 722049001

Sekretaris Sidang,



Zainul Islam, A.Md.Pi

Penguji I



Farihin Lazim, M. Tr. T.
NIDN: 0711099201

Penguji II,



Achmad Baijuri, M.Kom.
NIDN: 0715078902

Mengetahui
Dekan,



Abd. Ghofur, M.Kom.
NIDN: 711088303

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : **Syamsuddin**
NPM : 2020502030
Prodi : S-1 Sistem Informasi
Fakultas : Fakultas Sains dan Teknologi

Menyatakan dengan sebenarnya, bahwa Tugas Skripsi ini secara keseluruhan adalah hasil penelitian atau karya saya sendiri kecuali pada bagian bagian yang dirujuk sebagai sumber referensi dan disebutkan dalam daftar Pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa Tugas Skripsi ini hasil plagiasi, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Situbondo, 08 Agustus 2024

Saya yang menyatakan



Syamsuddin

MOTTO

KEBERHASILAN BUKANLAH MILIK ORANG YANG PINTAR.

TAPI, KEBERHASILAN ADALAH

MILIK ORANG YANG MAU UNTUK BERUSAHA.



PERSEMBAHAN

Saya persembahkan Skripsi ini kepada orang – orang yang telah membantu saya dalam menjalani pembuatan karya ilmiah dan program Skripsi ini:

1. Kedua orang tua saya beserta kakak - kakak saya dan seluruh keluarga yang berada dirumah yang selalu mendo'akan, menyemangati dan men-support saya.
2. Sahabat-sahabat seperjuangan mahasiswa Sistem Informasi angkatan 2020 yang terhebat dan berada dibawah naungan almamater Universitas Ibrahimy yang selalu memberikan semangat kepada saya.
3. Kepada seluruh anak kamar Sunan Kalijogo No.09 (D.09) yang selalu memberikan saya semangat dan juga mendoakan saya.
4. Semua orang yang tidak bisa saya sebutkan satu persatu namanya yang telah mendukung saya sepenuhnya sehingga saya dapat menyelesaikan laporan ini dengan sebaik baiknya.

KATA PENGANTAR

Bismillahirrahmanirrahim

Alhamdulillah, saya ucapkan kehadiran Allah SWT. yang telah melimpahkan rahmat dan karunia-Nya kepada saya, sehingga saya dapat menyelesaikan Skripsi ini dengan segala upaya dan jerih payah.

Skripsi ini merupakan salah satu syarat untuk memenuhi kurikulum S1 jurusan Sistem Informasi (SI) di Universitas Ibrahimi (UNIB) Sukorejo Situbondo Jawa Timur.

Penyusunan Skripsi ini tentunya tak luput dari bantuan berbagai pihak, maka dalam kesempatan ini saya akan ungkapkan terima kasih kepada:

1. KHR. Ach. Azaim Ibrahimi selaku Pengasuh Pondok Pesantren Salafiyah Syafi'iyah Sukorejo Situbondo.
2. KH. Ach. Fadhail, S.H., M.H. Selaku Rektor Universitas Ibrahimi Sukorejo Situbondo.
3. Abd. Ghofur, M. Kom. selaku dekan Fakultas Sains dan Teknologi Universitas Ibrahimi
4. Dr. Ach. Khumaidi, M.P. Selaku Wakil Dekan I Fakultas Sains dan Teknologi Universitas Ibrahimi
5. Abd. Wafi, M.P. Selaku Wakil Dekan II Fakultas Sains dan Teknologi Universitas Ibrahimi
6. Ahmad Lutfi, M. Kom. Selaku Wakil Dekan III Fakultas Sains dan Teknologi Universitas Ibrahimi

7. Achmad Bajuri, M. Kom. selaku Ketua Program Studi Sistem Informasi.
8. Muhammad Ali Ridla, M.Kom. selaku dosen pembimbing I dan Akhlis Munazilin, S. Kom., M.T Selaku dosen pembimbing II yang telah membantu kami dalam menyusun Tugas Skripsi ini.
9. Kasubag Ubudiyah di Pondok Pesantren Salafiyah Syafi'iyah Sukorejo selaku pihak yang telah memberikan kesempatan pada kami untuk meneliti bagian Ubudiyah tersebut.
10. Seluruh Dosen Fakultas Sains dan Teknologi yang telah memberikan kami ilmu sehingga sampai pada masa Skripsi ini

Penulis sadar bahwa tidak ada manusia yang luput dari salah dan lupa karena didunia tidak ada manusia yang sempurna, begitu pula dengan penyusunan Tugas Skripsi ini. Untuk itu penulis mengharapkan kritik dan saran dari para pembaca dalam penulisan selanjutnya sehingga berguna khususnya bagi penulis dan bagi pembaca pada umumnya.

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN.....	v
MOTTO.....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABLE.....	xiv
ABSTRAK.....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	3
1.4 Batasan Masalah.....	3
1.5 Tujuan Penelitian.....	3
1.6 Manfaat Penelitian.....	4
1.7 Metodologi Penelitian.....	4
1.7.1 Lokasi penelitian.....	5

1.7.2 Teknik Pengumpulan Data.....	5
1.8 Metode Penelitian.....	6
1.9 Sistematika Penulisan.....	6
BAB II LANDASAN TEORI	8
2.1 Teori yang Terkait Secara Umum	8
2.1.1 Data.....	8
2.1.2 Data Mining	12
2.1.3 Knowledge Discovery In Database (KDD)	21
2.1.4 Association Rules	24
2.1.5 Algoritma Apriori	25
2.1.6 Penelitian Terdahulu	28
BAB III ANALISIS DAN PERANCANGAN	31
3.1 Metode Penelitian.....	31
3.2 Bahan Penelitian.....	34
3.3 Metode pengumpulan data	34
3.3.1 Data Primer.....	34
3.3.2 Data Sekunder.....	35
3.4 Tahapan penelitian.....	36
3.5 Pengolahan Data Mining	37
3.5.1 Preprocessing	38
3.5.2 Transformation.....	39
3.5.3 Data Mining	40

3.5.4 Interpretation/Evaluasi.....	41
BAB IV	42
HASIL DAN PEMBAHASAN.....	42
4.1 Gambaran Umum Objek Penelitian.....	42
4.2 Hasil.....	42
4.3 Pembahasan	43
4.3.1 Data Transaksi Pelanggaran.....	43
4.3.2 Data Transaksi Tranformasi.....	45
4.3.3 Data Transaksi Dalam Bentuk Tabular.....	47
4.3.4 Pembentukan Itemset.....	48
4.3.5 Pembentukan Association Rules.....	53
BAB V.....	55
PENUTUP.....	55
5.1 Kesimpulan.....	55
5.2 Saran.....	55
DAFTAR PUSTAKA	57
LAMPIRAN.....	59
A. Surat Tugas Penelitian.....	59
B. Surat Keterangan Sudah Meneliti	60
BIODATA.....	61

DAFTAR GAMBAR

Gambar 2. 1 Tahapan-tahapan KDD.....	23
Gambar 2. 2 Ilustrasi Algoritma Apriori.....	28
Gambar 3. 1 Tahapan Penelitian	36
Gambar 4.1 Tabular Transaksi.....	47



DAFTAR TABLE

Tabel 3. 1 Transaksi Hasil Transformasi	39
Tabel 4. 1 Data Transaksi.....	43
Tabel 4. 2 Data Transaksi Hasil Transformasi.....	45
Tabel 4. 3 Pembentukan 1 Itemset	48
Tabel 4. 4 Hasil Pembentukan 1 Itemset.....	49
Tabel 4. 5 Pembentukan 2 Itemset	50
Tabel 4. 6 Hasil Pembentukan 2 Itemset.....	51
tabel 4. 7 Pembentukan 3 Itemset	52
tabel 4. 8 Hasil Pembentukan 3 Itemset	52
tabel 4. 9 Pembentukan 4 Itemset	53

ABSTRAK

Syamsuddin. 2024 **Penentuan Pola Pelanggaran Santri Pada Bagian Ubudiyah Menggunakan Algoritma Apriori (studi kasus: Pondok Pesantren Salafiyah Syafi'iyah Sukorejo.)** Skripsi, Program Studi Sistem Informasi, Universitas Ibrahimi. Pembimbing ; (1) Muhammad Ali Ridla, M.Kom, (2) Akhlis Munazilin, S.Kom., M.T

Data mining adalah proses penggalian pola dan informasi yang bermanfaat dari sejumlah besar data. Proses ini telah digunakan dalam berbagai hal, salah satunya adalah untuk menemukan pola pelanggaran santri di bagian ubudiyah. Pihak ubudiyah dapat menggunakan data mining dan algoritma apriori untuk mengolah data pelanggaran santri di bagian ubudiyah. algoritma yang digunakan untuk menentukan pola pelanggaran santri di bagian ubudiyah. Untuk menentukan pola kombinasi itemset dan aturan asosiasi, algoritma apriori adalah salah satu jenis aturan asosiasi, yaitu tidak berjama'ah, tidak Jum'at, tidak baca al-kahfi, yang memiliki nilai support 20% dan nilai confidence 38% kemudian hasil tidak berjama'ah, tidak shof, dan tidak jubah dengan nilai support 40% dan nilai confidence 75%.

Kata kunci: Data Mining, Algoritma Apriori, Aturan Asosiasi, Pelanggaran Ubudiyah.

ABSTRACT

Syamsuddin. 2024 Determining Patterns of Santri Violations in the Ubudiyah Section Using the Apriori Algorithm (case study: Salafiyah Syafi'iyah Sukorejo Islamic Boarding School.) Thesis, Information Systems Study Program, Ibrahimi University. Mentor; (1) Muhammad Ali Ridla, M.Kom, (2) Akhlis Munazilin, S.Kom., M.T

Data mining is the process of extracting patterns and useful information from large amounts of data. This process has been used in various ways, one of which is to find patterns of student violations in the Ubudiyah section. The ubudiyah can use data mining and a priori algorithms to process data on student violations in the ubudiyah section. algorithm used to determine patterns of student violations in the Ubudiyah section. To determine the combination pattern of itemsets and association rules, the a priori algorithm is one type of association rule, namely not congregational, not Friday, not reading al-kahfi, which has a support value of 20% and a confidence value of 38% then the results are not congregational. 'ah, no shof, and no cloak with a support value of 40% and a confidence value of 75%.

Keywords: Data Mining, Apriori Algorithm, Association Rules, Ubudiyah Violation

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data mining merupakan proses penggalian informasi dan pola bermanfaat dari data yang sangat besar. Data mining mencakup pengumpulan data, ekstraksi data, analisis data, dan statistic data. Data-data yang diolah dengan menggunakan Teknik data mining juga mampu menghasilkan pengetahuan yang sesuai dengan harapan. Menemukan pola yang sampai sekarang tidak teridentifikasi adalah tujuan dari penambangan data. Seharusnya tren Setelah diperoleh, ini dapat digunakan untuk mengatasi berbagai masalah. [1]

Kehidupan santri identik dengan kesederhanaan, disiplin, dan pengabdian diri kepada agama. Salah satu aspek penting dalam kehidupan santri adalah ubudiyah, yang mengacu pada kegiatan ibadah dan pengamalan agama Islam. Ubudiyah di pesantren memiliki peran sentral dalam membentuk karakter dan kepribadian santri. Santri dididik untuk menjalankan ibadah dengan khushyuk dan penuh makna, dan menerapkan prinsip-prinsip Islam dalam kehidupan sehari-hari. Penekanan pada Ibadah Wajib Santri diwajibkan untuk menjalankan ibadah wajib dengan sempurna dan tepat waktu.

Solusi yang dapat digunakan adalah meningkatkan pendidikan dan pembinaan intensif mengenai pentingnya ubudiyah. Melalui ceramah, kajian rutin, dan motivasi, Sistem pengawasan yang ketat juga sangat penting untuk memastikan kepatuhan santri terhadap aturan ubudiyah. Pengawasan ini bisa dilakukan melalui monitoring harian oleh pengurus pesantren atau ketua kamar,

Pengawasan yang efektif dapat mencegah terjadinya pelanggaran dan memberikan respon langsung kepada santri. Menunjukkan bahwa pengawasan yang sistematis dan berkelanjutan dapat meningkatkan disiplin santri secara signifikan.[2]

Teknik asosiasi dalam proses *Data Mining*. Aturan Asosiasi, juga disebut sebagai *Association Rules*, adalah suatu teknik untuk menentukan hubungan atau hubungan antara suatu benda dengan benda lainnya. Algoritma Apriori dapat digunakan dalam proses penemuan Association Rule. *Algoritma Apriori* adalah salah satu metode yang umum digunakan dalam *Data Mining* untuk menemukan Aturan Asosiasi dan kumpulan item yang sering muncul di database transaksional. Algoritma ini bekerja dengan mengidentifikasi setiap item yang ada dalam basis data dan menggabungkan kumpulan item yang lebih besar dengan syarat item tersebut muncul dengan frekuensi yang cukup tinggi dalam basis data.

Metode Apriori adalah sebuah teknik pembelajaran mesin dengan aturan, yang digunakan untuk menemukan keterkaitan korelasi menarik antara berbagai variabel dalam database besar. Tujuannya adalah menemukan pedoman ketat dalam database, memanfaatkan sejumlah tindakan yang relevan. Pendekatan ini merupakan teknologi untuk eksplorasi yang sangat fleksibel dan terutama digunakan untuk menemukan koneksi tersembunyi di antara data, sehingga menghasilkan pengelompokan dan klasifikasi berdasarkan berbagai tingkat detailnya. Metode Apriori ini menghasilkan aturan asosiasi dengan nilai kepercayaan (*confidence*) atau tingkat kepastian yang diukur.[3]

1.2 Identifikasi Masalah

Berdasarkan informasi latar belakang yang diberikan, ditemukan permasalahan dalam penelitian ini, yaitu Banyaknya pelanggaran yang terjadi pada bagian ubudiyah oleh santri, namun belum ada pola yang jelas yang bisa diidentifikasi untuk mengambil tindakan pencegahan.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan sebelumnya, rumusan masalah Penelitian ini mendemonstrasikan penggunaan Algoritma Apriori untuk menemukan pola pelanggaran santri di bagian ubudiyah.

1.4 Batasan Masalah

Harus ada batasan masalah mengingat masalah yang telah diidentifikasi sebelumnya yang dapat dilihat sebagai berikut:

- a. Data yang akan dianalisis adalah data pelanggaran santri di bagian ubudiyah pada bulan Desember 2023.
- b. Mengimplementasikan Algoritma Apriori untuk menemukan pola pelanggaran santri di bagian ubudiyah.
- c. Penelitian ini menggunakan perhitungan microsoft Excel untuk menghitung Algoritma Apriori.

1.5 Tujuan Penelitian

Tujuan utama penelitian ini adalah untuk lebih memahami pola perilaku santri dalam hal pelanggaran. yang mungkin terjadi di Bagian Ubudiyah menggunakan algoritma Apriori. Dengan menganalisis data pelanggaran santri secara sistematis, penelitian ini bertujuan untuk mengidentifikasi pola-pola yang

signifikan dalam perilaku pelanggaran, termasuk hubungan antara jenis pelanggaran, waktu pelanggaran, dan identitas santri yang terlibat. Pemahaman yang lebih dalam tentang pola-pola ini diharapkan dapat memberikan informasi yang mendalam kepada para pengelola pesantren untuk membantu mereka dalam membuat langkah-langkah yang lebih efektif untuk mencegah pelanggaran di masa mendatang.

1.6 Manfaat Penelitian

Penelitian mengenai penentuan pola pelanggaran santri di Bagian Ubudiyah menggunakan Algoritma Apriori memiliki beberapa manfaat yang signifikan. Pertama, penelitian ini dapat memberikan manfaat langsung bagi pengelola pesantren dengan memberikan pemahaman yang lebih baik tentang pola pelanggaran yang terjadi di lingkungan pesantren. Dengan mengetahui pola yang muncul, pengelola pesantren dapat menerapkan prosedur disiplin dan pengawasan dengan cara yang lebih sesuai dan efisien.

Selain itu, penelitian ini dapat membantu dalam meningkatkan efisiensi dan efektivitas pengawasan santri. Dengan menggunakan algoritma Apriori untuk menganalisis data pelanggaran, pengelola pesantren mampu menemukan pola yang mungkin terlewatkan oleh manusia. Hal ini memungkinkan mereka untuk mengalokasikan sumber daya pengawasan secara lebih cerdas, fokus pada area-area yang cenderung memiliki tingkat pelanggaran yang lebih tinggi.

1.7 Metodologi Penelitian

Metodologi penelitian merupakan proses dalam pelaksanaan penelitian skripsi ini. Dalam hal ini, Penulis menggunakan teknik ketika melakukan

penelitiannya. Berikut ini teknik yang digunakan oleh penulis dalam melakukan penelitian.

1.7.1 Lokasi penelitian

Dimana penelitian dan pengumpulan data dilakukan di Sub Bagian Ubudiyah Pondok Pesantren Salafiyah Syafi'iyah Sukorejo.

1.7.2 Teknik Pengumpulan Data

Teknik pengumpulan data yang dipakai dalam penelitian yang akan dilakukan untuk menyusun skripsi ini yaitu:

1. Observasi

Observasi langsung, dokumentasi, dan pengumpulan data dilakukan pada Sub Bagian Ubudiyah Pondok Pesantren Salafiyah Syafi'iyah Sukorejo.

2. Wawancara

Dengan cara berkomunikasi secara langsung dengan Pengurus Sub Bagian Ubudiyah Pondok Pesantren Salafiyah Syafi'iyah Sukorejo.

3. Studi Pustaka

Untuk berkontribusi pada pengetahuan tentang *literature* terkait penelitian, penulis menambahkan teknik pengumpulan data menggunakan studi kepustakaan dan informasi diperoleh dari literatur yang berkaitan dengan penulisan, serta pencarian data dan informasi dari surat kabar, tabloid, majalah maupun *internet* sebagai bahan referensi

1.8 Metode Analisis

Metode yang cocok digunakan untuk penelitian *Data Mining* apriori adalah kuantitatif. Karena *Algoritma Apriori* berfokus pada analisis data transaksional yang dapat diukur dan dihitung dalam bentuk angka, sehingga lebih sesuai dengan pendekatan kuantitatif dalam analisis data.

1.9 Sistematika Penulisan

Agar struktur penulisan lebih mudah dipahami dan dipatuhi skripsi ini, Peneliti menggunakan sistematika, atau fase, sebagai struktur dan pedoman penulisan, serta menguraikan Tahapan kegiatan selaras dengan ruang lingkup yang telah ditentukan sebelumnya. Skripsi ini terbagi menjadi beberapa bab, yang dijelaskan sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini membahas dan mengurai tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab ini, akan dibahas teori-teori ilmiah yang menjadi landasan utama bagi masalah yang diteliti. Bab ini terdiri dari dua bagian: teori dasar dan teori umum, yang mencakup konsep-konsep fundamental yang relevan dengan penelitian, dan teori-teori khusus yang berhubungan langsung dengan subjek penelitian.

BAB III : METODE PENELITIAN

Pada bab ini, akan dijelaskan tentang metode penelitian yang mencakup pengumpulan dan analisis data yang sesuai dengan tahapan KDD (*Knowledge Discovery in Databases*).

BAB IV : HASIL DAN PEMBAHASAN

Bagian ini akan memberikan pemahaman yang luas tentang subjek penelitian, langkah interpretasi dan penilaian, dan hasil dari proses Data Mining yang dilakukan dengan Algoritma Apriori. Proses ini menghasilkan aturan yang dapat digunakan untuk mengidentifikasi jenis pelanggaran santri pada aspek ubudiyah.

BAB V : PENUTUP

Bab ini berisi kesimpulan dari awal hingga akhir penelitian yang dilakukan oleh peneliti, serta menyajikan kritik dan saran yang membangun terhadap penelitian tersebut.

BAB II

LANDASAN TEORI

2.1 Teori yang Terkait Secara Umum

2.1.1 Data

Pengertian data secara umum adalah sekumpulan keterangan atau fakta mentah berupa simbol, angka, kata-kata, atau citra, yang didapatkan melalui proses pengamatan atau pencarian ke sumber-sumber tertentu.

Menurut bahasa, kata data berasal dari bahasa latin, yaitu datum yang berarti sesuatu yang diberikan. Data juga bisa diartikan sebagai hasil pengukuran atau pengamatan suatu variabel yang bentuknya dapat berupa simbol, warna, kata-kata, angka, atau citra.

Data adalah kumpulan fakta atau informasi yang dapat digunakan untuk menganalisis, memahami, atau memprediksi sesuatu. Data dapat berupa angka, kata, atau gambar. Data dapat diambil dari banyak sumber, seperti survei, eksperimen, atau catatan. Data dapat digunakan untuk membuat keputusan, memecahkan masalah, atau meningkatkan kinerja. Ini adalah beberapa contoh data:

- a. Hasil pelanggaran selama satu tahun
- b. Usia, jenis kelamin, dan lokasi pelanggaran
- c. Hasil uji coba sangsi baru
- d. Data cuaca selama satu bulan

Data dapat digunakan untuk berbagai tujuan, seperti:

- a. Menganalisis pelanggaran santri
- b. Menentukan kebutuhan santri
- c. Meningkatkan efisiensi operasional
- d. Membuat keputusan yang lebih baik

Data adalah aset yang berharga bagi bisnis, organisasi, dan individu. Dengan menggunakan data, didapatkan keputusan yang lebih baik, meningkatkan kinerja, dan mencapai tujuan yang lebih besar.

Menurut Al-Bahra, data adalah komponen informasi dasar yang kemudian akan diproses untuk menghasilkan informasi. Namun, Longkutoy menyatakan dalam bukunya "Pengenalan komputer" bahwa Istilah "data" merujuk pada informasi atau komponennya yang memiliki makna yang terkait dengan kenyataan. Simbol-simbol, gambar, angka, atau huruf dapat menjadi komponen data, atau elemen lain yang merepresentasikan ide, objek, kondisi, situasi, dan lain sebagainya [4].

Namun, beberapa ahli menyatakan dalam bukunya "Pengenalan komputer" bahwa data adalah istilah majemuk yang merujuk pada fakta atau bagian fakta yang memiliki makna yang terkait dengan kenyataan. Simbol-simbol, gambar, angka, atau huruf dapat menjadi komponen data.

Pengertian Data Menurut Para Ahli

Berikut akan dibahas apa saja definisi data menurut pendapat para ahli, dari dalam negeri dan luar negeri.

Menurut John J. Longkutoy (2012)

Data adalah suatu istilah majemuk yang berarti fakta atau bagian dari fakta yang mengandung arti yang dihubungkan dengan kenyataan, simbol-simbol, gambar-gambar, angka-angka, huruf-huruf, atau simbol-simbol yang menunjukkan suatu ide, objek, kondisi atau situasi, dan lain-lain.

Menurut Robert N. Antony dan John Dearden

Definisi data diartikan sebagai sebuah bentuk jamak dari data item atau bentuk tunggal datum.

Menurut Arikunto Suharsimi (2002)

Menurut Arikunto Suharsimi, pengertian data adalah semua fakta dan angka-angka yang dapat dijadikan bahan untuk menyusun sebuah informasi.

Menurut Siyoto dan Sodik (2015)

Pengertian data adalah sesuatu yang dikumpulkan oleh peneliti berupa fakta empiris yang digunakan untuk memecahkan masalah atau menjawab pertanyaan penelitian.

Menurut Nuzulla Agustina

Definisi data adalah keterangan mengenai suatu hal yang sudah sering terjadi dan berupa himpunan fakta, angka, grafik tabel, gambar, lambang, kata, huruf, yang menyatakan sesuatu pemikiran, objek, serta kondisi dan situasi.

Menurut Anhar

Arti data menurut Anhar adalah sesuatu yang diketahui atau dianggap oleh orang yang mengamatinnya.

Menurut Soeratno dan Lincoln Arsyad (1993)

Menurut Soeratno dan Arsyad, selain untuk memecahkan masalah, data juga perlu diadakan dalam rangka menguji suatu hipotesis yang berdasar pada suatu model. Wujud data dapat berbentuk sebagai angka, huruf, gambar, suara, suatu keadaan, atau simbol-simbol lainnya.

Menurut Sutabri (2012)

Definisi data adalah kenyataan yang menggambarkan suatu kejadian serta merupakan suatu bentuk yang masih mentah yang belum dapat bercerita banyak sehingga perlu diolah lebih lanjut melalui suatu model untuk menghasilkan informasi.

Menurut Kuswadi dan E. Mutiara

Pengertian data adalah sekumpulan keterangan yang diperoleh dari suatu pengamatan yang dapat berupa angka, lambang, atau sifat.

Menurut Zulkifli A.M.

Definisi data menurut Zulkifli A.M. merupakan suatu keterangan, bukti atau fakta tentang suatu kenyataan yang masih mentah (original) yang belum diolah sama sekali.

Menurut Gordon B. Davies

Data merupakan bahan mentah bagi informasi yang dirumuskan sebagai kelompok lambang-lambang tidak acak yang menunjukkan tindakan-tindakan, hal-hal dan sebagainya.

2.1.2 Data Mining

Hermawati mendefinisikan *Data Mining* sebagai proses yang menggunakan lebih dari satu metode pembelajaran komputer (machine learning) untuk secara otomatis menganalisis dan mengambil data [5]. *Data Mining* merupakan upaya untuk menemukan pola atau model baru yang valid, menguntungkan, dan mudah dipahami dalam database yang sangat besar (massive database). Tujuan dari *Data Mining* adalah mencari pola atau trend yang diinginkan dalam database besar untuk membantu pengambilan keputusan di masa mendatang. Pola-pola ini diidentifikasi oleh perangkat khusus yang memberikan analisis data yang berguna dan informatif, yang kemudian dapat dipelajari lebih lanjut secara detail.

Selain itu, *Data Mining* juga disebut sebagai metode untuk mengidentifikasi hubungan yang bermakna antara pola dan kecenderungan dengan menganalisis sekumpulan data yang sangat besar dalam penyimpanan menggunakan teknik pengenalan pola seperti matematika dan statistik [6].

Menurut Gartner Group, *Data Mining is the process of discovering new meaningful correlations, patterns and trends by sifting through large amounts of data stored in a repository, using pattern recognition technology as well as statistical and mathematical techniques.*

Selain definisi di atas, terdapat beberapa pengertian dan definisi data mining adalah sebagai berikut:

- a. *Data Mining* merupakan proses analisis dari sekumpulan data untuk mengidentifikasi hubungan yang tidak terduga dan menampilkan data

dengan cara yang berbeda dari yang dilakukan sebelumnya, sehingga informasi yang dihasilkan dapat dimengerti dan bermanfaat bagi pemilik data.

- b. *Data Mining* merupakan bidang yang menggabungkan metode pembelajaran mesin, pengenalan pola, statistik, basis data, dan gambar untuk mengatasi masalah mengumpulkan data dari database yang besar.
- c. *Data Mining* merupakan kombinasi dari statistik, kecerdasan buatan, dan penelitian basis data yang terus berkembang dan berkembang pesat.

Menurut Baskoro, tujuan dari adanya *Data Mining* dapat dibagi menjadi tiga kategori, yaitu: Akurat artinya informasi harus menjelaskan keadaan dengan sebenar-benarnya[7].

- a. Explanatory, tujuan ini bertujuan untuk memberikan penjelasan terhadap beberapa kegiatan observasi atau kondisi yang terjadi. Dalam *explanatory Data Mining*, kita mencari hubungan sebab-akibat dan mencoba memahami alasan di balik fenomena atau pola yang ditemukan dalam data.
- b. Confirmatory, tujuan ini adalah untuk mengkonfirmasi atau menguji suatu hipotesis yang telah ada sebelumnya. Dengan menggunakan *confirmatory Data Mining*, kita berusaha untuk membuktikan atau menegaskan kebenaran dari suatu teori atau asumsi yang sebelumnya telah diajukan.
- c. Exploratory, tujuan ini adalah untuk melakukan analisis pada data baru dengan tujuan menemukan hubungan atau pola yang mengejutkan atau tidak terduga. Dalam *Exploratory Data Mining*, kita secara aktif mencari informasi baru atau relasi yang sebelumnya tidak diketahui atau tidak

terduga, sehingga dapat memberikan wawasan baru dan pemahaman yang lebih mendalam terhadap data.

Sebelum memulai proses *Data Mining*, disarankan untuk memiliki pemahaman yang mendalam tentang potensi dan kemampuan data mining. Hal ini penting agar kegiatan *Data Mining* yang dilakukan sesuai dengan kebutuhan dan dapat menghasilkan informasi yang sebelumnya tidak diketahui, bersifat inovatif, dan memberikan manfaat bagi pengguna [8].

Data Mining pada dasarnya memiliki fungsi dan tugas untuk mengidentifikasi pola yang harus diungkapkan dalam prosesnya. Secara umum, tugas *Data Mining* dapat dikelompokkan menjadi dua kategori utama, yaitu:

- a. Tugas prediktif bertujuan untuk menghitung nilai atribut lainnya untuk memprediksi nilai atribut tertentu. Target atau variable tak bebas adalah istilah yang digunakan untuk menggambarkan atribut yang diprediksi. Di sisi lain, Variable bebas adalah atribut yang digunakan untuk membuat prediksi.
- b. Deskriptif, tugas deskriptif bertujuan untuk mengungkapkan pola-pola seperti korelasi, tren, cluster, trayektori, dan anomali dalam data. Proses *Data mining* deskriptif membutuhkan teknik, dan sering disebut sebagai penelitian. post-processing untuk validasi dan penjelasan hasilnya.

Pengelompokan *Data Mining* menurut kusrini [9], Berdasarkan jenis pekerjaan yang dapat dilakukan, dapat dibagi menjadi berbagai kategori, yaitu:

- a. Deskripsi

Tugas deskripsi berfokus pada mencari metode untuk menunjukkan pola dan tren yang tersembunyi dalam data. Misalnya, petugas pengumpulan suara dapat mencari cara untuk menjelaskan pola pemilih yang kurang profesional dalam kampanye presiden. Deskripsi pola dan tren sering memberikan penjelasan yang mungkin terhadap fenomena yang diamati.

b. Estimasi

Kategorisasi dan estimasi sama, tetapi variabel yang dimaksud adalah nilai numerik daripada kategori. Misalnya, Anda dapat menghitung tekanan sistolik pasien yang dirawat di rumah sakit berdasarkan umur, jenis kelamin, dan indeks berat badan mereka, dan tingkat natrium darah mereka. Metode *Data Mining* yang terdapat dalam estimasi antara lain *Linear Regression*, *Neural Network*, dan *Support Vector*.

c. Prediksi

Seperti estimasi dan klasifikasi; prediksi menunjukkan hal-hal yang mungkin atau belum terjadi di masa depan. Misalnya, untuk memprediksi harga beras selama tiga bulan ke depan, Metode *Data Mining* yang terdapat dalam prediksi antara lain *Linear Regression*, *Neural Network*, dan *Support Vector*.

d. Klasifikasi

Tujuan klasifikasi adalah untuk menempatkan variabel target ke dalam berbagai kategori. Misalnya, mengkategorikan pendapatan menjadi tiga kategori: tinggi, sedang dan rendah. Teknik data mining yang termasuk dalam kategori antara lain *Naive Bayes*, *K-Nearest*, *C4.5*, *ID3*, *CART*, *Linear Discriminant Analysis*, dan *Logistic Regression*.

e. Clustering

Clustering berfokus pada mengelompokkan catatan, pengamatan, atau kasus dalam kelompok yang sebanding. Sebuah *Cluster* adalah kumpulan *Record* yang mirip satu sama lain dan berbeda dari *Record* dalam *Cluster* lainnya. Misalnya, Tujuan audit akuntansi adalah untuk membagi perilaku finansial menjadi kelompok-kelompok dan menemukan perilaku mencurigakan. *Metode Data Mining* yang terdapat dalam *Clustering* antara lain *K-Means*, *K-Medoids*, *Self-Organizing Map (SOM)*, dan *Fuzzy C-Means*.

f. Asosiasi

Fokus tugas asosiasi adalah menemukan hubungan antara berbagai peristiwa yang terjadi pada satu waktu. Pendekatan asosiasi ini sering digunakan dalam analisis keranjang komoditas. Metode data mining yang bekerja sama termasuk *FP-Growth*, *A Priori*, *Coefficient of Correlation*, dan *Chi Square*.

Dengan memahami pengelompokan dan tugas-tugas *Data Mining* ini, pengguna dapat memilih metode dan pendekatan yang tepat untuk memenuhi tujuan analisis dan mengungkapkan wawasan yang berharga dari data yang dimiliki.

a. Data Selection

Tahap ini mencakup pilih himpunan data, membuat himpunan data yang dimaksud, atau berkonsentrasi pada subset variabel (sampel data) di mana penemuan (discovery) akan dilakukan. Keputusan disimpan dalam berkas yang berbeda dari basis data operasional.

b. Pre-Processing dan Pembersihan Data

Pre-Processing dan Pembersihan data bertujuan untuk menghilangkan suara dan data yang tidak konsisten, duplikat dan memperbaiki kesalahan data. Anda juga dapat menggabungkan data dengan data luar yang relevan.

c. Transformasi

Tahap ini melibatkan transformasi atau penggabungan data menjadi bentuk yang lebih sesuai untuk proses penambangan, termasuk agregasi.

d. Data Mining

Data Mining merupakan proses pencarian pola atau informasi yang menarik dari data yang dipilih sesuai dengan tujuan umum proses KDD.

e. Interpretasi / Evaluasi

Tahap ini mencakup interpretasi dari pola-pola yang dihasilkan dari *Data Mining*. Dilakukan evaluasi untuk mengevaluasi apakah informasi atau pola yang ditemukan sesuai atau bertentangan dengan fakta atau hipotesis sebelumnya. Informasi yang diperoleh dari pola-pola ini dipresentasikan dalam bentuk visualisasi.

Karakteristik Data Mining

Setelah tahu tujuannya, Anda sebaiknya juga mengenal karakteristik data mining sebelum benar-benar menjalankannya, yaitu:

1. Membutuhkan Banyak Data

Sebelum ada internet, proses pengumpulan data dilakukan lewat survei kepada konsumen secara langsung. Tentunya, cara itu membutuhkan banyak waktu dan tenaga.

Saat ini, semua aktivitas konsumen secara online menghasilkan data yang melimpah. Salah satunya, adalah berbagai kebiasaan konsumen yang tercatat di akun di media sosial atau data diri yang tersimpan pada saat membuat akun di marketplace atau toko online.

Semakin bertambahnya data di dunia online, semakin besar juga kebutuhan data untuk melakukan data mining dengan baik.

2. Data Tidak Selalu Akurat

Meskipun mendapatkan banyak data kian mudah, tidak semua data memiliki kualitas yang baik untuk mendukung keputusan bisnis yang tepat. Inilah salah satu karakteristik data mining yang patut diperhatikan.

Kenapa hal itu bisa terjadi? Alasannya, bisa saja data yang diberikan oleh konsumen di berbagai platform tidak lengkap atau bahkan tidak valid. Palsunya, belum semua orang merasa nyaman membagikan data pribadi mereka secara online.

Selain tidak ingin privasi mereka terganggu, beberapa orang juga khawatir data mereka digunakan untuk tujuan negatif atau takut menjadi korban kejahatan online. Penggunaan cookies pada website menjadi salah satu strategi untuk mengumpulkan data pengunjung secara otomatis. Namun, pengunjung website tetap bisa menolak datanya dikumpulkan dengan cara tersebut.

3. Dilakukan dengan Cara Manual dan Bantuan Tools

Data scientist bisa saja melakukan data mining sendiri jika data yang dikumpulkan dan diolah belum banyak. Namun, kalau jumlahnya semakin berlipat, penambahan data secara manual tentu kurang efektif, bukan?

Itulah mengapa pada proses pengumpulan data, machine learning menjadi hal yang akan sangat membantu, baik dari tool maupun algoritmanya. Meskipun ketika berlanjut pada proses evaluasi data, peran data scientist makin dibutuhkan untuk mendapatkan hasil lebih akurat.

Manfaat Data Mining

Ada banyak manfaat yang bisa didapatkan dari data mining, terutama dalam dunia bisnis, seperti:

1. Meningkatkan Kualitas Layanan

Meskipun memiliki produk yang berkualitas, tapi kalau bisnis tidak mampu memberikan layanan kepada konsumen dengan baik, pasti mereka akan kecewa,

Nah, data mining menjanjikan manfaat untuk bisa meningkatkan kualitas layanan bisnis. Sebagai contoh, dengan sebuah tool CRM, Anda jadi tahu kendala yang sering mereka hadapi dan menyediakan proses yang lebih baik agar kendala tersebut tidak terjadi pada konsumen lagi

2. Melakukan Analisis Risiko

Data mining juga cukup efektif dalam melakukan analisis risiko. Itulah kenapa upaya ini kerap dilakukan oleh institusi keuangan. Salah satu contohnya, untuk memproses nasabah bank yang mengajukan pinjaman.

Pada praktiknya, bank akan mengumpulkan riwayat keuangan nasabah dari berbagai bank dan penyedia layanan finansial lainnya. Nah, jika orang tersebut memiliki catatan yang buruk, bank akan menanggukhan permintaan pinjamannya.

3. Mengoptimalkan Pengeluaran

Manfaat data mining yang lain adalah mendeteksi pemborosan dan dan mengupayakan belanja perusahaan lebih optimal. Upaya ini cukup sering dilakukan di berbagai jenis industri.

Sebagai contoh, Anda memasarkan produk lewat iklan media sosial. Setelah melakukan data mining, ternyata sumber transaksi dari blog lebih banyak dibanding dari iklan yang dipasang.

Artinya, iklan di media sosial Anda kurang efektif. Dengan begitu, Anda bisa mengevaluasi kampanye iklan Anda dan mengalokasikan anggarannya untuk memaksimalkan blog. Misalnya, membeli tema premium untuk fitur lebih lengkap, atau mencari penulis blog berkualitas agar bisa membuat konten lebih rutin.

4. Bisa Lebih Memahami Konsumen

Kebutuhan konsumen akan suatu produk bisa saja berubah-ubah. Bisa karena tren, atau lainnya. Nah, agar produk atau layanan Anda senantiasa laris, pastikan penawaran produk tetap relevan dengan kebutuhan itu.

Dengan data mining, Anda bisa mencatat perubahan tersebut sehingga dapat memberikan solusi produk yang tepat.

Contohnya, saat Hari Raya, terdapat beberapa produk yang sering dibeli sesuai data mining yang dilakukan. Maka, Anda bisa menampilkan produk tersebut di halaman utama toko online. Dengan begitu, akan memudahkan konsumen dalam berbelanja produk.

5. Memastikan Produksi Berjalan Baik

Untuk memastikan produksi berjalan dengan baik, memanfaatkan data mining adalah langkah jitu. Data mining yang dilakukan bisa berupa mengumpulkan data penggunaan, kapasitas produksi, dan lainnya.

Dengan semua data yang dimiliki, Anda bisa tahu kapan harus melakukan pemeliharaan untuk menghindari perbaikan yang tidak terjadwal dan mengurangi produktivitas.

6. Mengelola Pasokan Bahan Baku Lebih Baik

Manfaat data mining selanjutnya adalah untuk mengetahui kapan permintaan pasar meningkat. Jadi, mereka bisa menyiapkan jumlah pasokan untuk menghadapi permintaan yang membludak. Dengan menyiapkan bahan baku jauh-jauh hari, efektivitas kerja tentu lebih baik,

2.1.3 Knowledge Discovery In Database (KDD)

Knowledge Discovery in Database (KDD) adalah proses mencari informasi berharga dan mengidentifikasi pola yang terdapat dalam data. Informasi ini tersembunyi dalam basis data yang berukuran besar, sebelumnya tidak diketahui, namun memiliki potensi untuk memberikan manfaat yang

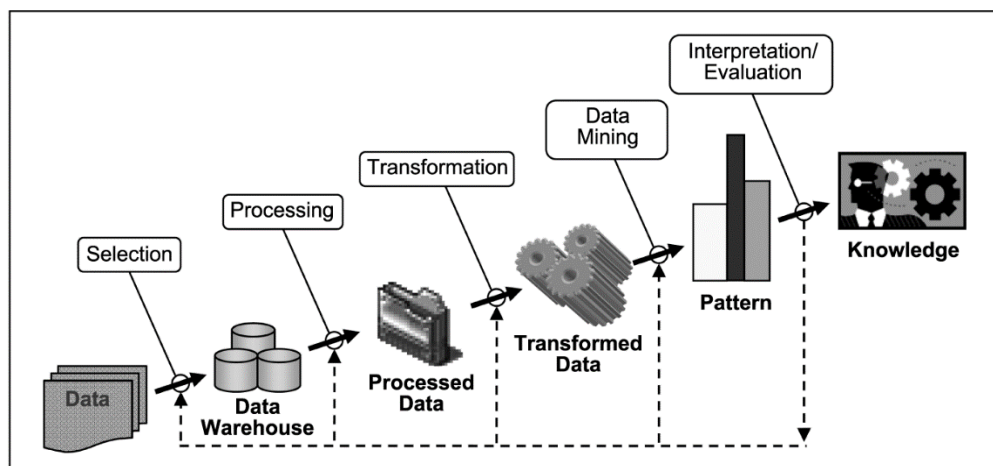
berarti. *Data Mining* merupakan salah satu tahap penting dalam rangkaian proses iteratif KDD.

Knowledge Discovery in Databases (KDD) atau Penemuan Pengetahuan dalam Basis Data adalah proses mengidentifikasi pola dan informasi yang berguna dari kumpulan data yang besar. KDD merupakan bidang interdisipliner yang menggabungkan teknik dari ilmu komputer, statistik, pembelajaran mesin, dan data mining untuk mengolah dan menganalisis data secara efektif. Proses KDD bertujuan untuk mengekstraksi pengetahuan yang berharga dan bisa ditindaklanjuti dari data, yang dapat digunakan untuk pengambilan keputusan, peramalan, dan optimisasi.

Tahapan utama dalam proses KDD meliputi seleksi data, pembersihan data, transformasi data, data mining, dan evaluasi pola. Seleksi data bertujuan memilih data yang relevan dari kumpulan data yang lebih besar. Pembersihan data adalah proses menghapus data yang tidak konsisten, tidak lengkap, atau salah. Transformasi data mengubah data mentah menjadi format yang lebih sesuai untuk analisis. Data mining merupakan inti dari KDD, di mana teknik algoritma digunakan untuk menemukan pola tersembunyi atau aturan dalam data. Akhirnya, tahap evaluasi pola digunakan untuk menilai dan menafsirkan hasil temuan untuk menentukan apakah pola tersebut berguna.

KDD memiliki aplikasi yang luas di berbagai bidang, seperti bisnis, kesehatan, ilmu sosial, dan teknologi informasi. Dalam bisnis, KDD dapat digunakan untuk menganalisis perilaku konsumen dan meningkatkan strategi

pemasaran. Di bidang kesehatan, KDD membantu dalam menemukan pola penyakit atau efektivitas pengobatan. Sementara itu, dalam ilmu sosial, KDD digunakan untuk menganalisis data survei dan penelitian untuk menemukan pola perilaku manusia. Secara keseluruhan, KDD merupakan alat yang sangat penting dalam era data besar untuk mengungkap wawasan yang tidak terlihat dari data yang kompleks dan besar



Gambar 2. 1 Tahapan-tahapan KDD

Tahapan Proses KDD terdiri dari:

a. Data Selection

Tahap ini mencakup pemilihan himpunan data, pembentukan himpunan data target, atau fokus pada subset variabel (sampel data) di mana penemuan (discovery) akan dilakukan. Hasil seleksi disimpan dalam berkas terpisah dari basis data operasional.

b. Pre-Processing dan Pembersihan Data

Pre-Processing dan Pembersihan Data bertujuan untuk menghapus data yang tidak konsisten dan menghilangkan noise, duplikasi data, serta

memperbaiki kesalahan data. Selain itu, data juga dapat diperkaya dengan data eksternal yang relevan.

c. Transformasi

Tahap ini melibatkan transformasi atau penggabungan data menjadi bentuk yang lebih tepat untuk proses mining, termasuk peringkasan (agregasi).

d. Data Mining

Data Mining merupakan proses pencarian pola atau informasi menarik dalam data terpilih dengan menggunakan teknik, metode, atau algoritma tertentu sesuai dengan tujuan dari keseluruhan proses KDD.

e. Interpretasi / Evaluasi

Tahap ini mencakup interpretasi dari pola-pola yang dihasilkan dari *Data Mining*. Dilakukan evaluasi untuk menguji apakah pola atau informasi yang ditemukan sesuai atau bertentangan dengan fakta atau hipotesis sebelumnya. Pengetahuan yang diperoleh dari pola-pola yang terbentuk dipresentasikan dalam bentuk visualisasi.

2.1.4 Association Rules

Association rule adalah metode yang digunakan untuk mengidentifikasi pola-pola yang sering muncul di antara berbagai transaksi, di mana setiap transaksi terdiri dari beberapa item. Tujuan utama dari Metode ini bertujuan untuk mendukung sistem rekomendasi dengan menemukan pola antar item dalam transaksi. Dengan demikian, metode ini membantu dalam mengidentifikasi asosiasi atau korelasi yang kuat antara barang atau barang yang dibeli secara bersamaan oleh pelanggan, yang dapat digunakan untuk

memberikan rekomendasi produk kepada pengguna berdasarkan kebiasaan pembelian mereka pelanggan[10].

Metodologi dasar analisis asosiasi terdiri dari dua tahap:

- a. Tahap analisis pola frekuensi tinggi.

Pada tahap ini, dilakukan pencarian kombinasi item yang memenuhi syarat *Minimum* dari nilai *Support* dalam database. Nilai *Support* sebuah item dihitung menggunakan rumus berikut:

$$\text{Support}(A) = \frac{\text{Jumlah_Pelanggaran_A}}{\text{Jumlah_Pelanggaran}}$$

Sedangkan untuk menghitung nilai *Support* dari 2 item, digunakan rumus berikut:

$$\text{Support}(A, B) = P(A \cap B) \frac{\text{Jumlah_Pelanggaran_A_dan_B}}{\text{Jumlah_Pelanggaran}}$$

- b. Tahap pembentukan aturan asosiatif.

Tahap berikutnya dimulai setelah penemuan semua pola frekuensi tinggi adalah mencari aturan asosiasi yang memenuhi syarat *Minimum* untuk *Confidence*. *Confidence* dari sebuah aturan $A \cup B$ dihitung dengan menggunakan rumus berikut:

$$\text{Confidence}(A, B) = \frac{\text{Pelanggaran_A} \cap B}{\text{Pelanggaran_A}}$$

2.1.5 Algoritma Apriori

Algoritma Apriori adalah algoritma dasar yang dikembangkan oleh Agrawal & Srikant pada tahun 1994 dengan tujuan untuk menemukan item-

item sering dalam aturan asosiasi Boolean. *Algoritma Apriori* termasuk dalam jenis teknik Aturan Asosiasi yang digunakan dalam *Data Mining*. Aturan yang menyatakan hubungan antara berbagai karakteristik biasanya disebut sebagai sebutan *Affinity Analysis* atau *Market Basket Analysis*. Analisis asosiasi atau *Association Rule* mining adalah teknik *Data Mining* yang digunakan untuk menentukan aturan untuk gabungan item yang muncul bersamaan dalam data.

Analisis pola adalah fase analisis asosiasi yang menarik perhatian banyak peneliti frekuensi tinggi atau *Frequent Pattern Mining*. Para peneliti tertarik untuk menghasilkan algoritma yang efisien dalam tahap ini, karena pentingnya dalam mengidentifikasi pola-pola yang sering muncul.

Pentingnya suatu asosiasi dapat diukur melalui dua parameter, yaitu *Support* dan *confidence*. *Support* (nilai penunjang) mencerminkan persentase jumlah item yang digabungkan dalam database, sedangkan *confidence* (nilai kepastian) menggambarkan seberapa kuat hubungan antara aturan asosiasi tersebut [9].

Algoritma Apriori mengikuti sejumlah tahap yang dikenal sebagai alur atau pass [11].

- a. Tahap pembentukan kandidat itemset.

Kandidat k-itemset dibuat dari kombinasi (k-1)-itemset yang diperoleh dari iterasi sebelumnya. Salah satu metode dalam *Algoritma Apriori* adalah mengurangi kandidat k-itemset di mana subsetnya berisi k-1 pola frekuensi tinggi yang tinggi tidak termasuk k-1.

- b. Tahap penghitungan *Support* dari tiap kandidat k-itemset.

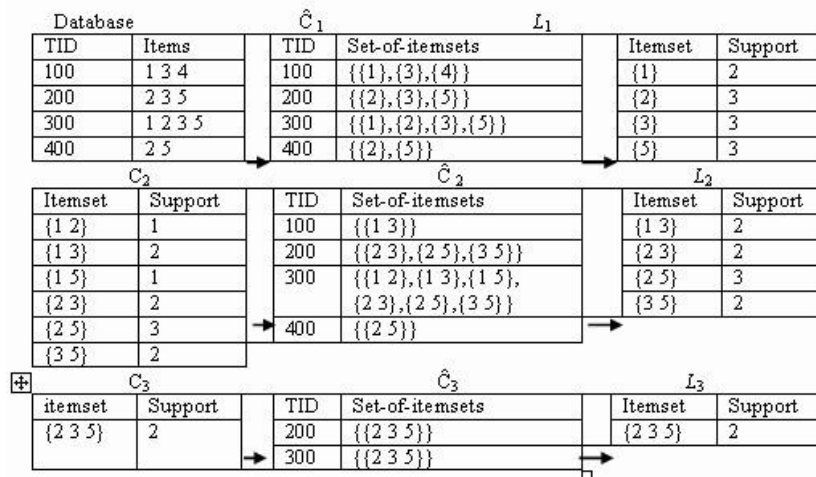
Support dari setiap kandidat k-itemset dihitung dengan memindai database untuk menghitung total transaksi yang mencakup setiap item dalam kandidat k-itemset tersebut. Ini adalah karakteristik utama dari *Algoritma Apriori*, di mana perlu dilakukan perhitungan berulang sebanyak k-itemset terpanjang dalam seluruh database.

c. Menetapkan pola frekuensi tinggi.

Pola frekuensi tinggi yang mencakup k-item atau k-itemset ditentukan dari kandidat k-itemset yang memiliki *Support* lebih besar dari *Minimum Support* yang ditetapkan sebelumnya.

d. Jika tidak ada pola frekuensi tinggi baru ditemukan, maka seluruh proses dihentikan.

Untuk memperoleh pemahaman yang lebih mendalam tentang proses *Algoritma Apriori*, akan diberikan sebuah contoh penggunaan *Algoritma Apriori* dengan mengacu pada database pada gambar 2.2 dan asumsi *Minimum Support* sebesar 2 transaksi.



Gambar 2. 2 Ilustrasi Algoritma Apriori

2.1.6 Penelitian Terdahulu

“ Penerapan Data Mining Untuk Menentukan Pola Penjualan Sepatu Menggunakan Metode Algoritma Apriori ”

Perusahaan PT Global Vision, merupakan sebuah perusahaan supplier sepatu yang berlokasi didaerah Jakarta Barat. Didalam meningkatkan kinerja karyawan dan profit perusahaannya,

PT Global Vision telah menerapkan penggunaan teknologi yang sedang trend saat ini, tetapi untuk proses perancangan strategi bisnis pemasarannya perusahaan ini belum memiliki acuan khusus dan masih mengandalkan dari perkiraan saja. Selain itu perusahaan ini belum cukup mampu dalam mengelola pertumbuhan data-data mereka yang sekian lama semakin menumpuk dalam jumlah yang besar, sehingga mengakibatkan data-data tersebut tidak menghasilkan informasi yang dapat bernilai lebih, kemudian data-data tersebut menjadi tidak berarti serta tidak berguna. Untuk menindaklanjuti dari

permasalahan tersebut, maka penelitian ini dibuat dengan tujuan agar perusahaan PT Global Vision mendapatkan sebuah acuan ataupun pedoman khusus didalam penganalisaan strategi pemasaran produk sepatu mereka guna menunjang keberhasilan dari strategi pemasaran bisnisnya. Penelitian ini menggunakan algoritma apriori, dimana algoritma tersebut mampu melakukan pencarian frequent itemset dan telah banyak digunakan pada data transaksi.

Kemudian hasil dari penelitian ini, PT Global Vision dapat melakukan perancangan strategi bisnisnya dengan baik, dimana perusahaan ini mampu memprediksi dalam mempersiapkan persediaan stok merk sepatu pada periode selanjutnya. [12]

“ Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat ”

Data mining merupakan proses untuk mendapatkan informasi yang berguna dari gudang basis data yang berupa ilmu pengetahuan. penelitian ini melakukan analisa data dengan menggunakan data mining dan metode

algoritma apriori. Sistem yang dibangun ditujukan untuk pemenuhan dalam penentuan pola pembelian obat dengan menggunakan bahasa pemrograman Visual Basic 6.0 dan database Mysql pada studi kasus di sektor kesehatan. Sistem ini dibangun berdasarkan kebutuhan pengguna yang diperoleh melalui metode wawancara dan studi lapangan. Metodologi pengembangan sistem yang digunakan yaitu metode waterfall yang terdiri Analisis, Desain, Pengkodean dan Pengujian. Hasil pengujian dengan algoritma apriori dan sistem yang dibangun menunjukkan hasil yang telah

memenuhi kebutuhan dalam penentuan pola pembelian obat berdasarkan kecenderungan pembelian obat oleh pelanggan. Dibandingkan dengan sistem yang sedang berjalan kinerja tersebut ditunjukkan pada efektifitas informasi dari sistem tentang penentuan pola pembelian obat untuk ketersediaan obat dan tata letak obat untuk memudahkan dalam mengetahui keberadaan obat yang dilihat dari 2 itemset obat.[13]

” Penerapan Data Mining dengan Metode Algoritma Apriori untuk Menentukan Pola Pembelian Ikan ”

Tujuan pembuatan aplikasi data mining menggunakan metode Algoritma apriori adalah untuk menentukan pola pembelian jenis ikan yang paling diminati pada UD. Mumu Jaya. UD. Mumu Jaya Pandeglang adalah usaha dagang yang bergerak di bidang penjualan ikan laut di pasar pandeglang. UD. Mumu Jaya merupakan suatu jenis usaha penjualan ikan laut untuk menentukan ikan yang akan dibeli untuk persediaan gudang. Saat ini semua kegiatan transaksi yang berhubungan dengan penjualan dan pembelian masih menggunakan buku catatan seperti untuk pendataan ikan, pendataan penjualan ikan, dan masih memperkirakan ikan yang akan dibeli untuk memenuhi persediaan ikan pada gudang, selain itu masih belum adanya model perhitungan untuk menentukan belanja persediaan ikan yang paling diminati oleh konsumen. Untuk mendukung kegiatan tersebut maka perlu dibuat aplikasi yang dapat membantu permasalahan tersebut, yakni dengan dibuatnya suatu aplikasi data mining menggunakan metode Algoritma apriori untuk menentukan jenis ikan yang paling diminati oleh konsumen. [14]

BAB III

ANALISIS DAN PERANCANGAN

3.1 Metode Penelitian

Metode yang digunakan untuk penelitian *Data Mining Apriori* adalah metode kuantitatif. *Algoritma Apriori* difokuskan pada analisis data transaksional yang dapat diukur dan dihitung dalam bentuk angka, sehingga lebih cocok untuk digunakan dalam pendekatan kuantitatif dalam analisis data.

Penggunaan metode kuantitatif dalam penelitian *Data Mining Apriori* memungkinkan pengguna untuk melakukan analisis statistik dan pengukuran yang lebih tepat terhadap pola dan asosiasi di dalam data transaksional. Dalam metode kuantitatif, kita dapat menggunakan berbagai teknik matematis dan statistik untuk mengidentifikasi dan menguji kekuatan hubungan antara data item transaksional.

Dengan bantuan pendekatan kuantitatif, dapat diukur seberapa sering suatu pola muncul dalam data, menghitung tingkat kepercayaan dan dukungan dari asosiasi antara item, serta mengidentifikasi pola yang paling relevan dan signifikan. Hal ini dapat membantu dalam pengambilan keputusan berdasarkan bukti yang kuat dan dapat diandalkan dari data.

Selain itu, metode kuantitatif juga memungkinkan kita untuk melakukan analisis *Data Mining* secara lebih efisien dan akurat, terutama ketika kita memiliki jumlah data transaksional yang besar. Dengan mengolah data dalam bentuk angka dan menerapkan teknik-teknik kuantitatif, kita dapat

menghindari bias dan interpretasi subyektif yang mungkin terjadi dalam analisis kualitatif.

Secara keseluruhan, penggunaan metode kuantitatif dalam penelitian *Data Mining Apriori* memberikan keuntungan dalam menggali pola dan asosiasi yang tidak terlihat dalam data transaksi secara lebih tepat, obyektif, dan efisien. Hal ini memungkinkan untuk mendapatkan wawasan yang lebih dalam dan informasi berharga dalam mengambil keputusan di bidang yang berbeda, seperti bisnis, pemasaran, ilmu sosial, dan lain sebagainya.

Karakteristik metode penelitian kuantitatif adalah sebagai berikut:

- 1) Pengumpulan data berbasis angka: Metode penelitian kuantitatif mengumpulkan data dalam bentuk angka atau data kuantitatif. Data ini biasanya terdiri dari angka-angka, ukuran, skor, atau variabel yang dapat diukur secara objektif.
- 2) Pendekatan deduktif: Metode kuantitatif menggunakan pendekatan deduktif dalam merumuskan hipotesis dan menguji teori. Hipotesis dihasilkan berdasarkan teori atau kerangka konseptual yang ada, dan kemudian data dikumpulkan untuk menguji hipotesis tersebut.
- 3) Representasi dan analisis statistik: Metode digunakan untuk menganalisis data yang telah dikumpulkan statistik untuk mengidentifikasi pola, hubungan, atau asosiasi yang mungkin ada di dalam data.
- 4) Objektivitas dan replicability: Metode kuantitatif menekankan pada objektivitas dalam pengumpulan dan analisis data. Hal ini berarti data dan

analisis harus dapat diulang oleh peneliti lain untuk memastikan keabsahan temuan.

- 5) Generalisasi: Hasil dari penelitian kuantitatif cenderung dapat digeneralisasi ke populasi yang lebih luas. Dalam metode kuantitatif, pengambilan sampel yang representatif digunakan untuk menghasilkan kesimpulan yang dapat mewakili populasi yang lebih besar.
- 6) Penggunaan instrumen penelitian: Metode kuantitatif sering menggunakan instrumen penelitian yang telah diuji keandalannya untuk mengumpulkan data. Contoh instrumen penelitian meliputi kuesioner, tes, atau perangkat pengukuran lainnya.
- 7) Tesis dan antitesis: Metode kuantitatif melibatkan perumusan hipotesis (tesis) dan hipotesis alternatif (antitesis) yang akan diuji melalui analisis data. Tujuannya adalah untuk mencari dukungan empiris terhadap salah satu dari kedua hipotesis tersebut.
- 8) Skala besar: Metode kuantitatif sering digunakan dalam penelitian dengan skala besar, di mana data yang dikumpulkan dari banyak partisipan atau unit observasi.
- 9) Ketepatan dan keakuratan: Metode kuantitatif cenderung memberikan tingkat ketepatan dan keakuratan yang tinggi dalam mengukur variabel-variabel tertentu dan dalam menguji hubungan antara variabel-variabel tersebut.

Dengan karakteristik-karakteristik tersebut, Penelitian kuantitatif mengumpulkan dan menganalisis angka, sehingga dapat memberikan

pemahaman yang kuat tentang fenomena yang diteliti dan memungkinkan pengujian teori secara empiris.

3.2 Bahan Penelitian

Bahan penelitian yang digunakan dalam penelitian ini kemudian diubah menjadi acuan meliputi data pelanggaran santri sebagai objek yang diteliti. Data yang dikumpulkan dari berbagai sumber tersebut akan menjadi dasar analisis untuk mengidentifikasi pola, dan asosiasi dalam transaksi pelanggaran, sehingga memberikan wawasan yang berharga dalam memahami perilaku santri dan mengoptimalkan agar pengurus di bidang ubudiyah bisa mengurangi terjadinya pelanggaran.

3.3 Metode pengumpulan data

3.3.1 Data Primer

Peneliti mengumpulkan data secara langsung melalui dua metode untuk mendapatkan data primer, yaitu wawancara (interview) dan observasi di bagian ubudiyah.

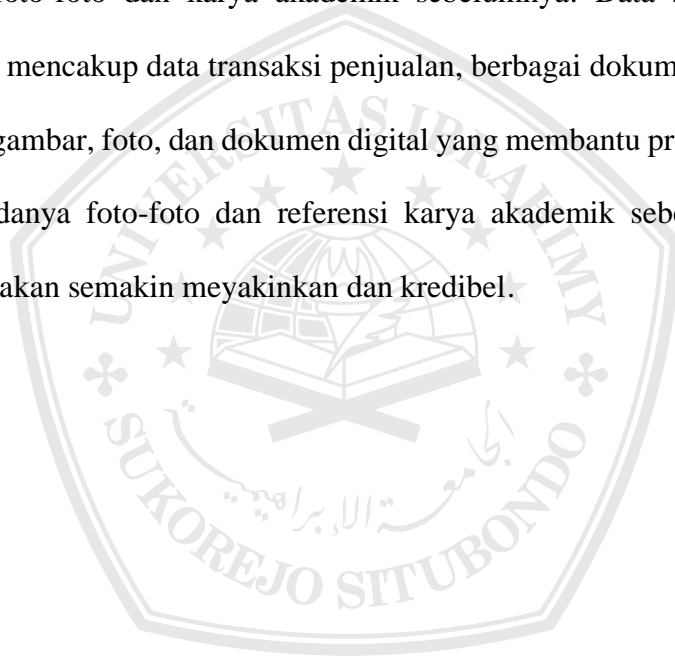
Pertama, metode wawancara (interview) digunakan untuk mendapatkan informasi tentang pelanggaran santri pada bagian ubudiyah. Data yang dihimpun dari wawancara ini akan menjadi bagian dari data yang dianalisis dalam perhitungan *Data Mining*. Fokus utama dalam analisis adalah data transaksi pelanggaran santri pada bulan Desember 2023.

Kedua, metode observasi dilakukan dengan melihat kejadian secara langsung pada perilaku santri tentang ke ubudiyahan. Observasi ini

memberikan data langsung tentang proses pelanggaran, interaksi santri dengan petugas, serta aspek-aspek lain yang relevan untuk penelitian.

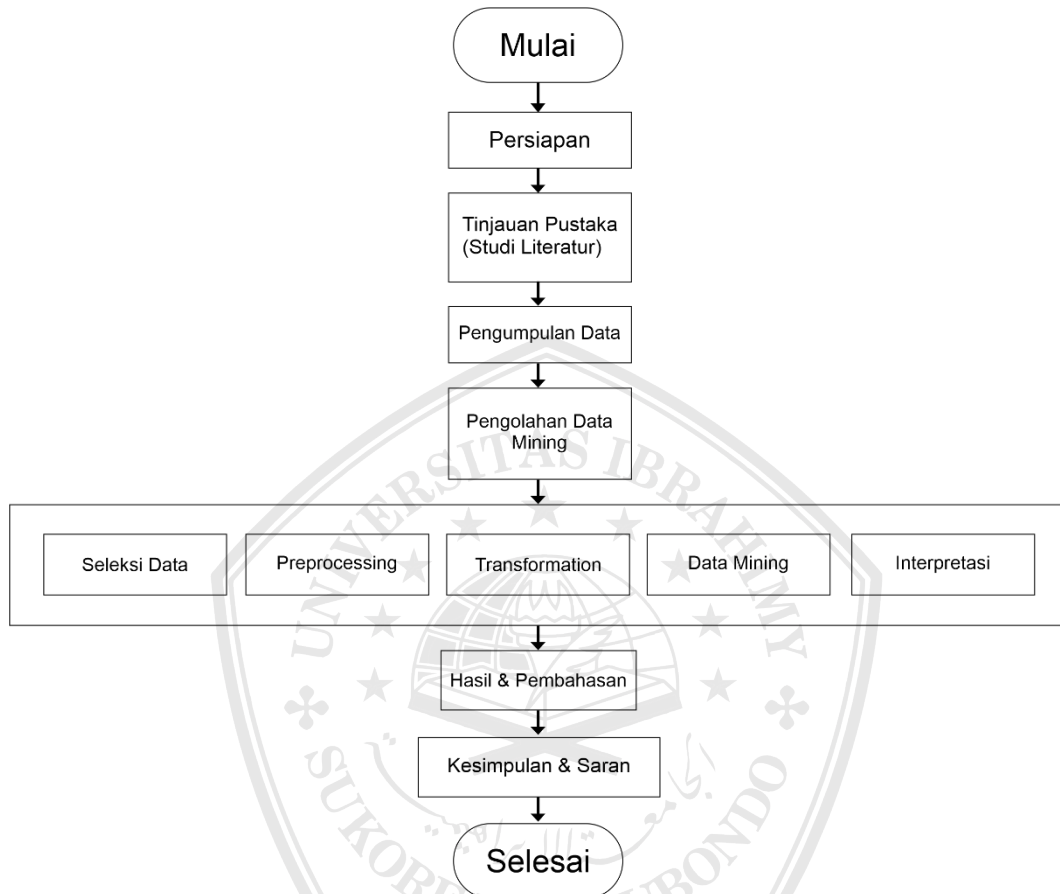
3.3.2 Data Sekunder

Dalam mendapatkan data sekunder, peneliti mengumpulkan berbagai data terkait dengan penelitian, khususnya data transaksi pelanggaran. Keberhasilan Penelitian ini akan menjadi lebih kredibel karena didukung oleh berbagai foto-foto dan karya akademik sebelumnya. Data sekunder yang digunakan mencakup data transaksi penjualan, berbagai dokumen tertulis atau softcopy, gambar, foto, dan dokumen digital yang membantu proses penulisan. Dengan adanya foto-foto dan referensi karya akademik sebelumnya, hasil penelitian akan semakin meyakinkan dan kredibel.



3.4 Tahapan penelitian

Gambar 3. 1 Tahapan Penelitian



Kerangka penelitian yang tergambar dalam diagram alir ini mencakup proses keseluruhan penelitian yang akan dilakukan. Penelitian ini akan melakukan hal-hal berikut:

- a. Tahap Persiapan: Pada tahap ini, peneliti akan mengidentifikasi subjek atau populasi dari petugas tim kontrol, dengan objek penelitian yang difokuskan pada data pelanggaran santri. Selain itu, batasan penelitian akan ditentukan dan rencana penelitian akan disusun.

- b. Tahap Tinjauan Kepustakaan: Pada tahap ini, peneliti akan melakukan telaah SLiMS perpustakaan dan studi literatur yang terkait dengan penjualan dan topik yang relevan.
- c. Tahap Pengumpulan Data: Data akan dikumpulkan melalui tiga cara, yaitu wawancara dengan petugas, observasi, dan dokumen terkait.
- d. Tahap Pengolahan *Data Mining*: Data yang terkumpul akan diolah menggunakan metode Knowledge Discovery in Database (KDD) untuk melakukan proses *Data Mining*.
- e. Tahap Hasil dan Pembahasan: Pada tahap ini, hasil dari proses *Data Mining* menggunakan *Algoritma Apriori* akan dijelaskan dan didiskusikan.
- f. Tahap Kesimpulan dan Saran: Pada langkah terakhir, peneliti akan menghasilkan kesimpulan berdasarkan hasil penelitian dan memberikan rekomendasi yang berkaitan dengan pihak petugas agar dapat meningkatkan kinerja dan efisiensi.

3.5 Pengolahan Data Mining

Pengolahan *Data Mining* penelitian ini dilakukan dengan mengikuti langkah-langkah yang telah ditetapkan dalam Knowledge Discovery in Database (KDD). Tujuannya adalah untuk menghasilkan data yang tepat untuk urutan yang telah direncanakan. Berikut adalah langkah-langkah dalam pengolahan *Data Mining* tersebut:

3.5.1 Preprocessing

Pada tahap ini, preprocessing akan melibatkan dua proses utama, yaitu integrasi data dan data cleaning, sebelum dapat melanjutkan ke tahap mining. Kedua prosedur tersebut dijelaskan di sini:

- a) Integrasi Data: Integrasi data adalah prosedur untuk menggabungkan data dari berbagai database, sehingga data tersebut saling terhubung dan berintegrasi. Pada tahapan ini, data dihubungkan dari tabel data penjualan yang berasal dari satu database, sehingga tidak ada penggabungan data yang diperlukan karena sumbernya berasal dari satu sumber yang sama.
- b) Data Cleaning: Data cleaning merupakan tahap awal dari proses Knowledge Discovery in Database (KDD). Pada tahap ini, data redundan, tidak relevan, dan tidak memiliki nilai harus dihapus dibersihkan. Tujuan dari data cleaning adalah untuk menghasilkan dataset yang bersih dan siap digunakan untuk tahap selanjutnya yaitu *Data Mining*. Data dikatakan tidak relevan jika tidak berkontribusi pada analisis yang ingin dilakukan. Missing value adalah ketika ada atribut dalam dataset yang tidak memiliki nilai atau kosong. Sedangkan data redundan adalah ketika dalam satu dataset terdapat lebih dari satu record yang berisi nilai yang sama. Setelah proses data cleaning dilakukan, dataset yang lebih sesuai dengan kriteria data penjualan akan dihasilkan.

Dengan melakukan integrasi data dan data cleaning pada tahap preprocessing, maka dataset yang sudah disiapkan akan lebih relevan, lengkap,

dan bersih sehingga dapat diandalkan dalam langkah selanjutnya, yaitu *Data Mining*.

3.5.2 Transformation

Tahap Transformation adalah langkah mengubah data agar sesuai dengan proses *Data Mining*.

Tabel 3. 1 Transaksi Hasil Transformasi

Nama	Jenis Pelanggaran
Nuzulul Farhan	T. Berjama'ah, T. Wirid, T. Shof, T. Jubah
Calvin Adi Hidayatullah	T. Berjama'ah, T. Tahajjud, T. Wirid, T. Munjyat, T. Shof, T. Jubah
Fathiz Zaman Ahmad	T. Jum'at, T. Baca Al-Kahfi
Ainur Rofiqi	T. Berjama'ah, T. Shof, T. Jubah
Kamilul Hasib	T. Tahajjud, T. Tarhim
Zainul Hasan	T. Tahajjud
Ubaidillah Fajri Ramadhani	T. Berjama'ah, T. Jum'at, T. Baca Al-Kahfi
Masdarul Hikam	T. Berjama'ah, T. Wirid, T. Munjyat, T. Shof, T. Jubah
Yafis Huki Fissiyani	T. Syawariq
Muhadzabul Irfan	T. Berjama'ah, T. Tahajjud, T. Shof, T. Jubah
Muhammad Walid Al-Alief	T. Baca Al-Kahfi
Muhammad Danish Ahza	T. Berjama'ah, T. Wirid, T. Munjyat, T. Shof, T. Jubah
Akhmad Fadillah	T. Tahajjud, T. Syawariq
Ahmad Sufyan	T. Tarhim, T. Syawariq
Muhammad Alif Furqon	T. Berjama'ah, T. Wirid, T. Munjyat, T. Shof, T. Jubah
Saiful Haq Zainullah	T, Jum'at, T. Baca Al-Kahfi
Ahmad Surya Pratama	T. Berjama'ah, T. Jubah

Muhammad Rifaldo F.	T. Berjama'ah, T. Shof, T. Jubah
Rama Danil Firdaus A.	T. Jum'at
Abdul Hannan	T. Berjama'ah, T. Wirid, T. Munjiyat, T. Shof, T. Jubah
Dzulfikar Maulana El Ghazi	T. Berjama'ah, T. Shof, T. Jubah
Najmi Bassam Zamani	T. Jum'at
Atha Shafy Abqory Altaf	T. Jum'at, T. Baca Al-Kahfi
Muh Irfan Mandar	T. Berjama'ah, T. Shof, T. Jubah
Robih Akmalul Fawaid	T. Syawariq
Bagus Prasetyo	T. Berjama'ah, T. Tahajjud, T. Shof, T. Jubah
Galang Ahmad Safariyanto	T. Jum'at, T. Baca Al-Kahfi
Moh. Malka Syarief Firdaus	T. Tarhim, T. Syawariq
Moh. Faqih Iklil Maulana	T. Berjama'ah, T. Shof, T. Jubah
Prarega Putra Fauzandy R.	T. Berjama'ah, T. Jubah

Berdasarkan tabel 3.1 Proses menggunakan dua atribut. *Data Mining*, yaitu atribut tanggal transaksi dan jenis pelanggaran.

3.5.3 Data Mining

Setelah melalui proses persiapan data dan penyesuaiannya untuk teknik *Data Mining*, langkah selanjutnya adalah melakukan proses *Data Mining*. Tahap ini bertujuan untuk menemukan pola atau data menarik dalam data yang telah dipilih, menggunakan metode *Association Rule* dan *Algoritma Apriori* yang sesuai dengan tujuan dari keseluruhan proses.

Data yang digunakan dalam proses ini merupakan data transaksi pelanggaran santri, bulan Desember 2023, dengan total 30 transaksi pelanggaran.

3.5.4 Interpretation/Evaluasi

Setelah mendapatkan pola penjualan dari proses *Data Mining*, langkah selanjutnya dalam proses Knowledge Discovery in Database (KDD) adalah tahap Interpretation/Evaluasi. Tahap Ini adalah komponen penting dari proses KDD yang melibatkan menentukan apakah pola atau data ditemukan sesuai dengan fakta atau teori sebelumnya.

Pola data yang dihasilkan dari data mining harus ditunjukkan sehingga pihak yang berkepentingan dapat memahaminya. Pola atau informasi dalam metode Rule of Association dibuat dalam bentuk aturan yang diperoleh dari perhitungan Algoritma Apriori. Dalam tahap Interpretation/Evaluasi, rules ini akan diuji dan dievaluasi kebenarannya berdasarkan data yang ada dan juga pengetahuan domain yang relevan.

Selama proses Interpretation/Evaluasi, dilakukan validasi untuk memastikan bahwa pola atau data yang diperoleh dari proses *Data Mining* memiliki nilai yang signifikan dan dapat memberikan wawasan yang berarti. Hasil evaluasi akan membantu dalam menentukan keputusan atau tindakan lebih lanjut berdasarkan temuan yang telah diidentifikasi dari proses KDD. Proses ini merupakan langkah kritis untuk memastikan bahwa hasil dari proses *Data Mining* dapat memberikan kontribusi yang berarti bagi tujuan akhir dari analisis tersebut

BAB IV

HASIL DAN PEMBAHASAN

4.1 Gambaran Umum Objek Penelitian

Sub. Bagian Ubudiyah berada dalam naungan Kabag. Kepesantrenan pada Bidang Kepesantrenan dan PU Pondok Pesantren Salafiyah Syafi'iyah Sukorejo. Sub Bag. Ubudiyah memiliki fungsi dalam keubudiyahan yang berhubungan dengan beribadahnya santri.

Sebagai komponen yang bertanggung jawab atas urusan Kepesantrenan di Pondok Pesantren Salafiyah Syafi'iyah Sukorejo. Sub. Bagian Ubudiyah di pimpin oleh Kasubag. Ubudiyah ustad Suhamar Iskandar, S.Ag. Yang berada dibawah Kabag. Kepesantrenan dan bertanggung jawab secara langsung kepada pengasuh dan secara teknis mendapat pembinaan dari Kabag. Kepesantrenan dan Kabid Kepesantrenan dan PU.

4.2 Hasil

Setelah menganalisis perancangan *Data Mining* yang bertujuan untuk mengidentifikasi pola pelanggaran santri pada bagian ubudiyah berdasarkan kategori atau jenis pelanggaran, menggunakan *Algoritma Apriori*, penelitian ini dapat menemukan pola pelanggaran berdasarkan kategori atau jenis pelanggaran. Untuk memudahkan proses *Data Mining*, peneliti menggunakan Microsoft Excel dengan data transaksi pelanggaran pada bulan desember 2023 yang diambil dari sub Bagian Ubudiyah. Hasil analisis *Data Mining* ini

menghasilkan informasi tentang pola pelanggaran yang dapat digunakan untuk mengambil keputusan lebih efektif dalam mengatasi pelanggaran santri.

4.3 Pembahasan

Pada bagian ini, akan dijelaskan langkah-langkah pelaksanaan implementasi *Data Mining* dengan menggunakan *Algoritma Apriori*, yang terdiri dari beberapa langkah berikut:

4.3.1 Data Transaksi Pelanggaran

Berdasarkan data transaksi pelanggaran pada bulan Desember 2023 Dapat dilihat pada tabel 4.1 dibawah ini

Tabel 4. 1 Data Transaksi

Nama	Item Pelanggaran
Nuzulul Farhan	1,4,8,10
Calvin Adi Hidayatullah	1, 3, 4, 5, 8, 10
Fathiz Zaman Ahmad	2, 6
Ainur Rofiqi	1, 8, 10
Kamilul Hasib	3, 7
Zainul Hasan	3
Ubaidillah Fajri Ramadhani	1, 2, 6
Masdarul Hikam	1, 4, 5, 8, 10
Yafis Huki Fissiyan	9
Muhadzabul Irfan	1, 3, 8, 10

Muhammad Walid Al-Alief	6
Muhammad Danish Ahza	1, 4, 5, 8, 10
Akhmad Fadillah	3, 9
Ahmad Sufyan	7, 9
Muhammad Alif Furqon	1, 4, 5, 8, 10
Saiful Haq Zainullah	2, 6
Ahmad Surya Pratama	1, 10
Muhammad Rifaldo F.	1, 8, 10
Rama Danil Firdaus A.	2
Abdul Hannan	1, 4, 5, 8, 10
Dzulfikar Maulana El Ghazi	1, 8, 10
Najmi Bassam Zamani	2
Atha Shafy Abqory Altaf	2, 6
Muh Irfan Mandar	1, 8, 10
Robih Akmalul Fawaid	9
Bagus Prasetyo	1, 3, 8, 10
Galang Ahmad Safariyanto	2, 6
Moh. Malka Syarief Firdaus	7, 9
Moh. Faqih Iklil Maulana	1, 8, 10
Prarega Putra Fauzandy R.	1, 10

No	Jenis Pelanggaran
1	Tidak Berjama'ah
2	Tidak Jum'at
3	Tidak Tahajjud
4	Tidak Wirid
5	Tidak Munjiyat
6	Tidak Baca Al-Kahfi
7	Tidak Tarhim
8	Tidak Shof
9	Tidak Syawariq
10	Tidak Jubah

4.3.2 Data Transaksi Tranformasi

Tabel 4. 2 Data Transaksi Hasil Transformasi

Transaksi	Jenis Pelanggaran
1	T. Berjama'ah, T. Wirid, T. Shof, T. Jubah
2	T. Berjama'ah, T. Tahajjud, T. Wirid, T. Munjiyat, T. Shof, T. Jubah
3	T. Jum'at, T. Baca Al-Kahfi
4	T. Berjama'ah, T. Shof, T. Jubah
5	T. Tahajjud, T. Tarhim
6	T. Tahajjud
7	T. Berjama'ah, T. Jum'at, T. Baca Al-Kahfi

8	T. Berjama'ah, T. Wirid, T. Munjiyat, T. Shof, T. Jubah
9	T. Syawariq
10	T. Berjama'ah, T. Tahajjud, T. Shof, T. Jubah
11	T. Baca Al-Kahfi
12	T. Berjama'ah, T. Wirid, T. Munjiyat, T. Shof, T. Jubah
13	T. Tahajjud, T. Syawariq
14	T. Tarhim, T. Syawariq
15	T. Berjama'ah, T. Wirid, T. Munjiyat, T. Shof, T. Jubah
16	T, Jum'at, T. Baca Al-Kahfi
17	T. Berjama'ah, T. Jubah
18	T. Berjama'ah, T. Shof, T. Jubah
19	T. Jum'at
20	T. Berjama'ah, T. Wirid, T. Munjiyat, T. Shof, T. Jubah
21	T. Berjama'ah, T. Shof, T. Jubah
22	T. Jum'at
23	T. Jum'at, T. Baca Al-Kahfi
24	T. Berjama'ah, T. Shof, T. Jubah
25	T. Syawariq
26	T. Berjama'ah, T. Tahajjud, T. Shof, T. Jubah
27	T. Jum'at, T. Baca Al-Kahfi
28	T. Tarhim, T. Syawariq
29	T. Berjama'ah, T. Shof, T. Jubah

30	T. Berjama'ah, T. Jubah
----	-------------------------

Tabel 4.2 memuat data tentang jumlah pelanggaran dilakukan oleh santri, dan data Ini akan digunakan sebagai dasar untuk melakukan perhitungan dengan menggunakan *Algoritma Apriori*.

4.3.3 Data Transaksi Dalam Bentuk Tabular

Data transaksi penjualan barang telah diatur dalam bentuk tabel, yang memudahkan dalam mengidentifikasi jumlah item yang terdapat dalam setiap transaksi, seperti yang ditunjukkan pada Gambar 4.1 di bawah ini.

Gambar 4.1 Tabular Transaksi

Transaksi	T. Berjama'ah	T. Jum'at	T. Tahajjud	T. Wirid	T. Munjiyat	T. Baca Al-Kahf	T. Tarhim	T. Shof	T. Syawarid	T. Jubah
Nuzulul Farhan	1	0	0	1	0	0	0	1	0	1
Calvin Adi Hidayatullah	1	0	1	1	1	0	0	1	0	1
Fathiz Zaman Ahmad	0	1	0	0	0	1	0	0	0	0
Ainur Rofiqi	1	0	0	0	0	0	0	1	0	1
Kamilul Hasib	0	0	1	0	0	0	1	0	0	0
Zainul Hasan	0	0	1	0	0	0	0	0	0	0
Ubaiddillah Fajri Ramadhani	1	1	0	0	0	1	0	0	0	0
Masdarul Hikam	1	0	0	1	1	0	0	1	0	1
Yafis Huki Fissiyani	0	0	0	0	0	0	0	0	1	0
Muhadzabul Irfan	1	0	1	0	0	0	0	1	0	1
Muhammad Walid Al-Atief	0	0	0	0	0	1	0	0	0	0
Muhammad Danish Ahza	1	0	0	1	1	0	0	1	0	1
Akhmad Fadillah	0	0	1	0	0	0	0	0	1	0
Ahmad Sufyan	0	0	0	0	0	0	1	0	1	0
Muhammad Alif Furqon	1	0	0	1	1	0	0	1	0	1
Saiful Haq Zainullah	0	1	0	0	0	1	0	0	0	0
Ahmad Surya Pratama	1	0	0	0	0	0	0	0	0	1
Muhammad Rifaldo F.	1	0	0	0	0	0	0	1	0	1
Rama Danil Firdaus A.	0	1	0	0	0	0	0	0	0	0
Abdul Hannan	1	0	0	1	1	0	0	1	0	1
Dzulfikar Maulana El Ghazi	1	0	0	0	0	0	0	1	0	1
Najmi Bassam Zamani	0	1	0	0	0	0	0	0	0	0
Atha Shafy Abqory Altaf	0	1	0	0	0	1	0	0	0	0
Muh Irfan Mandar	1	0	0	0	0	0	0	1	0	1
Robih Akmalul Fawaid	0	0	0	0	0	0	0	0	1	0
Bagus Prasetyo	1	0	1	0	0	0	0	1	0	1
Galang Ahmad Safariyanto	0	1	0	0	0	1	0	0	0	0
Moh. Malka Syarief Firdaus	0	0	0	0	0	0	1	0	1	0
Moh. Faqih Ikliil Maulana	1	0	0	0	0	0	0	1	0	1
Prarega Putra Fauzandy R.	1	0	0	0	0	0	0	0	0	1

4.3.4 Pembentukan Itemset

a) Pembentukan 1 Itemset

Untuk mengidentifikasi pembentukan 1 Itemset dengan menentukan jumlah *Minimum Support 20%*, Langkah pertama adalah menggunakan data yang tercantum dalam transaksi dalam Tabel 4.2 Selanjutnya, jumlah barang dalam setiap transaksi dihitung, dan data ini akan diproses menggunakan rumus berikut ini:

$$\text{Support}(A) = \frac{\text{Jumlah_Pelanggaran_A}}{\text{Total_Pelanggaran}}$$

Tabel 4. 3 Pembentukan 1 Itemset

No	Jenis Item	Total	Support(%)	Min Support = 20%
1	T. Berjama'ah	16	53%	LULUS
2	T. Jum'at	7	23%	LULUS
3	T. Tahajjud	6	20%	LULUS
4	T. Wirid	6	20%	LULUS
5	T. Munjiyat	5	17%	TIDAK LULUS
6	T. Baca Al-Kahfi	6	20%	LULUS
7	T. Tarhim	3	10%	TIDAK LULUS
8	T. Shof	13	43%	LULUS
9	T. Syawariq	5	17%	TIDAK LULUS
10	T. Jubah	15	50%	LULUS

Dari proses pembentukan Itemset pada tabel 4.3 dengan *Minimum Support 20%* dapat diketahui yang memenuhi standart *Minimum Support* yaitu sebagai berikut:

Tabel 4. 4 Hasil Pembentukan 1 Itemset

No	Kombinasi 1 Item	Total	Nilai Support (%)
1	T. Berjama'ah	16	53%
2	T. Jum'at	7	23%
3	T. Tahajjud	6	20%
4	T. Wirid	6	20%
5	T. Baca Al-Kahfi	6	20%
6	T. Shof	13	43%
7	T. Jubah	15	50%

b) Kombinasi 2 Itemset

Selanjutnya, penyelidikan hasil dilanjutkan pembentukan pola untuk dua Itemset. Hal ini diperoleh dari rumus berikut saat menciptakan item set sering untuk satu item:

$$\text{Support}(A \cap B) = \frac{\text{Jumlah Transaksi A dan B}}{\text{Total Pelanggaran}}$$

Hasil dari proses pembentukan 2 Itemset pelanggaran tercantum dalam tabel berikut:

Tabel 4. 5 Pembentukan 2 Itemset

No	Jenis Item	Total	Support(%)	Min Support = 20%
1	T. Berjama'ah, T. Jum'at	1	3%	TIDAK LULUS
2	T. Berjama'ah, T. Tahajjud	3	10%	TIDAK LULUS
3	T. Berjama'ah, T. Wirid	5	17%	TIDAK LULUS
4	T. Berjama'ah, T. Munjiyat	5	17%	TIDAK LULUS
5	T. Berjama'ah, T. Baca Al-Kahfi	1	3%	TIDAK LULUS
6	T. Berjama'ah, T. Shof	12	40%	LULUS
7	T. Berjama'ah, T. Jubah	14	47%	LULUS
8	T. Jum'at, T. Baca Al-Kahfi	6	20%	LULUS
9	T. Tahajjud, T. Wirid	1	3%	TIDAK LULUS
10	T. Tahajjud, T. Munjiyat	1	3%	TIDAK LULUS
11	T. Tahajjud, T. Tarhim	1	3%	TIDAK LULUS
12	T. Tahajjud, T. Shof	3	10%	TIDAK LULUS
13	T. Tahajjud, T. Syawariq	1	3%	TIDAK LULUS
14	T. Tahajjud, T. Jubah	3	10%	TIDAK LULUS
15	T. Wirid, T. Munjiyat	5	17%	TIDAK LULUS
16	T. Wirid, T. Shof	4	13%	TIDAK LULUS
17	T. Munjiyat, T. Shof	5	17%	TIDAK LULUS
18	T. Munjiyat, T. Jubah	5	17%	TIDAK LULUS
19	T. Tarhim, T. Syawariq	2	7%	TIDAK LULUS
20	T. Shof, T. Jubah	6	20%	LULUS

Setelah melakukan proses pembentukan 2 Itemset, ditentukan nilai *Minimum Support* sebesar 20%, Berikut adalah hasil dari, serupa dengan pola frekuensi pertama pembentukan 2 Itemset, sebagaimana ditunjukkan dalam tabel di bawah ini:

Tabel 4. 6 Hasil Pembentukan 2 Itemset

No	Kombinasi 2 Item	Total	Nilai Support (%)
1	T. Berjama'ah, T. Shof	12	40%
2	T. Berjama'ah, T. Jubah	14	47%
3	T. Jum'at, T. Baca Al-Kahfi	6	20%
4	T. Shof, T. Jubah	6	20%

c) Kombinasi 3 Itemset

Selanjutnya, dilanjutkan dengan mencari hasil pembentukan pola untuk tiga Itemset. Hal ini diperoleh dari rumus berikut saat menciptakan item set sering untuk satu item:

$$\text{Support}(A \cap B \cap C) = \frac{\text{Jumlah Transaksi A, B dan C}}{\text{Total Pelanggan}}$$

Hasil dari proses pembentukan 3 Itemset pelanggaran tercantum dalam tabel berikut:

tabel 4. 7 Pembentukan 3 Itemset

No	Jenis Item	Total	Support(%)	Min Support = 20%
1	T. Berjama'ah, T. Wirid, T. Munjiyat	5	17%	TIDAK LULUS
2	T. Berjama'ah, T. Jum'at, T. Baca Al-Kahfi	6	20%	LULUS
3	T. Berjama'ah, T. Shof, T. Jubah	12	40%	LULUS
4	T. Jum'at, T. Shof, T. Jubah	3	10%	TIDAK LULUS

Setelah melakukan proses pembentukan 3 Itemset, ditentukan nilai *Minimum Support* sebesar 20%, seperti pada pola frekuensi pertama. Berikut adalah hasil dari pembentukan 3 Itemset, sebagaimana tertera dalam tabel berikut :

tabel 4. 8 Hasil Pembentukan 3 Itemset

No	Kombinasi 3 Item	Total	Nilai Support (%)
1	. Berjama'ah, T. Jum'at, T. Baca Al-Kahfi	6	20%
2	T. Berjama'ah, T. Shof, T. Jubah	12	40%

d) Kombinasi 4 Itemset

Selanjutnya, dilanjutkan dengan mencari hasil pembentukan pola untuk empat Itemset. Hal ini diperoleh dari hasil menggunakan rumus berikut untuk membuat set item sering untuk satu item:

$$Support(A \cap B \cap C \cap D) = \frac{Jumlah_Transaksi_A_B_C\ dan_D}{Total_Pelanggaran}$$

Hasil dari proses pembentukan 4 Itemset pelanggaran tercantum dalam tabel berikut:

tabel 4. 9 Pembentukan 4 Itemset

No	Jenis Item	Total	Support(%)	Min Support = 20%
1	T. Berjama'ah, T. Wirid, T. Shof, T. Jubah	5	17%	TIDAK LULUS
2	T. Berjama'ah, T. Tahajjud, T. Shof, T. Jubah	3	10%	TIDAK LULUS

Setelah melakukan proses pembentukan 4 Itemset, ditentukan nilai *Minimum Support* sebesar 20%, Maka tidak ada yang memenuhi dari *Minimum Support* 20%.

4.3.5 Pembentukan Association Rules

Setelah menemukan semua pola dengan sukses Pembentukan Itemset, Langkah berikutnya adalah menciptakan *Association Rules*. Selain itu, aturan yang memenuhi persyaratan dicari *Minimum* untuk *Confidence* sebesar 30%, dengan aturan "Jika A maka B" dan "Jika B maka A" [15]. Setiap nilai *Support* dari itemset mungkin akan tetap sama, tetapi nilai *confidencenya* bisa berbeda.

Tujuannya adalah untuk mengetahui nilai *Confidence* terbesar dari setiap itemset [16]. Kombinasi item Pelanggaran A, B Dan C dapat diselesaikan dengan rumus sebagai berikut:

$$\text{Confidence} = P(C/B/A) = \frac{\text{Pelanggaran}_A \cap B \cap C}{\text{Pelanggaran}_A}$$

Maka untuk kombinasi item pelanggaran yang bersamaan adalah Tidak Berjama'ah, Tidak Jum'at, Tidak Baca Al-Kahfi Dan Tidak Berjama'ah Tidak Shof, Tidak Jubah, dan dapat dilihat pada tabel berikut:

No	Pembentukan Kombinasi Tiga Item	Support (%)	Confidence (%)
1	T. Berjama'ah, T. Jum'at T. Baca Al-Kahfi	20%	38%
2	T. Berjama'ah, T. Shof T. Jubah	40%	75%

BAB V

PENUTUP

5.1 Kesimpulan

Data Mining dengan *Metode Asosiasi Algoritma Apriori* sangat membantu dalam menentukan hubungan pelanggaran pada bagian ubudiyah yang paling sering dilakukan oleh santri. Informasi ini menjadi sangat berharga dalam mencegah dan mengevaluasi terjadinya pelanggaran ubudiyah serta membantu menentukan pola pelanggaran santri.

Menurut penelitian yang dilakukan menggunakan 30 data transaksi pelanggaran, Hasil pelanggaran yang paling sering adalah T. Berjama'ah (53%), T. Jum'at (23%), T. Baca Al-Kahfi (20%), T. Tahajjud (20%), T. Shof (43%), dan T. Jubah (50%).

Selain itu, ada banyak kombinasi pelanggaran yang didapatkan dari hasil penelitian, yaitu kombinasi T. Berjama'ah dan T. Shof dengan nilai *Support* (40%), T. Jum'at Dan T. Baca Al-Kahfi (20%). Data ini dapat digunakan untuk menentukan rencana mengatasi pelanggaran yang lebih efisien dan meningkatkan lebih disiplin dalam kegiatan keubudiyahan.

5.2 Saran

Untuk peneliti yang ingin menentukan pola pelanggaran santri pada bagian ubudiyah, disarankan untuk melakukan pendekatan yang holistik dan berbasis data. Peneliti harus mengkombinasikan berbagai metode seperti observasi langsung, wawancara dengan santri dan pengurus pondok, serta

analisis catatan pelanggaran yang terdokumentasi. Menggunakan alat analisis statistik dapat membantu dalam mengidentifikasi tren dan pola pelanggaran secara lebih akurat. Selain itu, penting untuk mempertimbangkan faktor kontekstual seperti lingkungan, jadwal kegiatan ibadah, dan latar belakang santri. Dengan pendekatan ini, peneliti dapat memberikan rekomendasi yang lebih spesifik dan aplikatif untuk memperbaiki sistem pembinaan ubudiyah di pondok pesantren.



DAFTAR PUSTAKA

- [1] M. N. Muhammad arhami, *Data Mining Algoritma Dan Implementasi*. Yogyakarta: Penerbit Andi, 2022.
- [2] Yusuf, “jurnal pendidikan dan pembinaan karakter santri,” *Pendidik. dan Pembina. Karakter*, 2018.
- [3] A. Gupta, S. Jain, dan A. Tiwari, “Optimization and Improvement of Association Rule Mining using Genetic Algorithm and Fuzzy Logic,” *SSRN Electron. J.*, hal. 2299–2304, 2019, doi: 10.2139/ssrn.3358761.
- [4] Al-Bahra bin Ladjamuddin B., *Konsep Sistem Basis Data dan Implementasinya*. Yogyakarta, Indonesia: Graha Ilmu, 2004.
- [5] Hermawati, F. Astuti, dan P. Christian, *Data Mining*.
- [6] D. T. Larose, *Discovering knowledge in data: An introduction to data mining*. New Jersey: Inc, 2005.
- [7] H. Baskoro, “Implementasi Algoritma K-Means Menggunakan Data Penyewaan Alat Berat Untuk Melakukan Estimasi Nilai Outcome,” Universitas Pembangunan Nasional “Veteran,” 2010.
- [8] E. Prasetyo, *Data Mining Mengolah Data Menjadi Informasi menggunakan Matlab*. Yogyakarta, Indonesia: Andi Offset, 2014.
- [9] Kusriani, Luthfi, dan E. Taufiq, *Algoritma Data Mining*. Yogyakarta, Indonesia: Andi Offset, 2009.

- [10] Fadlina, "Data Mining Untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori (Studi Kasus Di Polsekta Medan Sunggal)," *Inf. dan Teknol. Ilm.*, vol. 3, 2014.
- [11] D. Dinda Setiawati, "Penggunaan Metode Apriori untuk Analisa Keranjang Pasar pada Data Transaksi Penjualan Minimarket menggunakan Java dan MySql".
- [12] S. Syahriani, "Penerapan Data Mining Untuk Menentukan Pola Penjualan Sepatu Menggunakan Metode Algoritma Apriori," *Bina Insa. Ict J.*, vol. 9, no. 1, hal. 43, 2022, doi: 10.51211/biict.v9i1.1758.
- [13] R. Yanto dan R. Khoiriah, "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat," *Creat. Inf. Technol. J.*, vol. 2, no. 2, hal. 102, 2015, doi: 10.24076/citec.2015v2i2.41.
- [14] S. Saefudin dan S. DN, "Penerapan Data Mining Dengan Metode Algoritma Apriori Untuk Menentukan Pola Pembelian Ikan," *JSiI (Jurnal Sist. Informasi)*, vol. 6, no. 2, hal. 36, 2019, doi: 10.30656/jsii.v6i2.1587.
- [15] H. Santoso, I. P. Hariyadi, dan Prayitno, "Data Mining Analisa Pola Pembelian Produk," *Tek. Inform.*, no. 1, hal. 19–24, 2016.
- [16] A. R. Riszky dan M. Sadikin, "Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan Data Mining using Apriori Algorithm for Product Recommendation for Customers," vol. 7, no. December 2018, hal. 103–108, 2019, doi: 10.14710/jtsiskom.7.3.2019.103

LAMPIRAN

A. Surat Tugas Penelitian

**PANITIA TUGAS AKHIR DAN SKRIPSI**
UNIVERSITAS IBRAHIMY
FAKULTAS SAINS DAN TEKNOLOGI
JL. KHR. Syamsul Arifin No. 1-2 PO. Box 2 Phone (0852-3333-7581)
Fax. (0338) 453068 Situbondo 68374 website: www.ibrahimy.ac.id e-mail : umib2018@ibrahimy.ac.id
SUMBEREJO BANYUPUTIH SITUBONDO JAWA TIMUR

Nomor : 0828/405.91/071.095/M.3/VI/2024
Prihal : Permohonan Izin Penelitian

11 Juni 2024

Kepada
Yth. Kasubag. Ubudiyah Pondok Pesantren Salafiyah Syafi'iyah

Di Tempat

Assalamu 'alaikum warahmatullahi wabarakatuh.

Disampaikan dengan hormat, dalam rangka pelaksanaan penelitian guna penyusunan Skripsi. Judul **"Penentuan Pola Pelanggaran Santri Pada Bagian Ubudiyah Menggunakan Algoritma Apriori (Studi Kasus: Pondok Pesantren Salafiyah Syafi'iyah Sukorejo)"** dengan ini kami bermaksud mengajukan permohonan izin penelitian bagi mahasiswa kami :

Nama : SYAMSUDDIN
NIM : 2020502030
Program Studi : S1 Sistem Informasi
Waktu : 12 - 20 Juni 2024


Untuk melakukan penelitian di lembaga yang Bapak/Ibu pimpin. Sebagai bahan pertimbangan bersama ini kami lampirkan :

1. Proposal Skripsi
2. Fotocopy Kartu Tanda Mahasiswa

Demikian permohonan kami, atas berkenannya kami sampaikan *Jazakumullahu khairan.*

Wassalamu 'alaikum warahmatullahi wabarakatuh.

Ketua Panitia,



Dr. Ach. Khumaidi, M.P.

B. Surat Keterangan Sudah Meneliti**PONDOK PESANTREN "SALAFIYAH SYAFI'YAH" SUKOREJO
BIDANG KEPESANTRENAN DAN PU**

SUMBEREJO BANYUPUTIH SITUBONDO JAWA TIMUR
Akte Notaris No. 4/25.08.1970 & No. 3/05.07.2001

SURAT PERYATAAN**Nomor : 0828/318/B.a.2/VII/2024**

Yang bertanda tangan dibawah ini

Nama : **Suhamar Iskandar, S.Ag.**

Jabatan : Kasubag. Ubudiyah

Alamat : Sukorejo Banyuputih Situbondo

Menerangkan dengan sebenarnya bahwa

Nama : **SYAMSUDDIN**

NPM : 2020502030

Program Studi : S1 Sistem Informasi

Universitas : Universitas Ibrahimi Sukorejo Situbondo

**Mahasiswa tersebut benar benar telah melakukan Kegiatan Penelitian
di (Sub. Ubudiyah Putra) Pondok Pesantren Salafiyah Syafi'iyah Sukorejo.**

Demikian surat keterangan ini di buat, dengan sebenarnya dan dapat di pergunakan
sebagaimana semestinya.

Sukorejo, 20 Juni 2024

Kasubag Ubudiyah,



Suhamar Iskandar, S.Ag

BIODATA

Nama : SYAMSUDDIN
NPM : 2020502030
TTL : Situbondo, 08 Agustus 2002
Alamat Asal : Mojosari - Asembagus – Situbondo

Nama Orang Tua :

- Ayah : H. Ali Maburr
- Ibu : Hj. Musyarofah
- Alamat : Mojosari - Asembagus - Situbondo

Jenjang Pendidikan :

- 2009-2014 : MI Al-Azhar Mojosari
- 2014-2017 : MTs Al-Azhar Mojosari
- 2017-2020 : SMK Ibrahimy 1 Sukorejo
- 2020-2024 : Universitas Ibrahimy Situbondo