

HALAMAN MOTTO

“Ad Maiora Natus Sum.”
Aku dilahirkan untuk hal yang lebih besar



HALAMAN PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dengan penuh rasa syukur kepada Allah SWT,
karya sederhana ini kupersembahkan kepada:

1. Kedua Orang Tua tercinta yang dengan kasih sayang, doa, dan pengorbanannya tiada henti memberikan semangat dan kekuatan dalam setiap langkah hidupku.
2. Mas Huda, Mba Uni, Mas Baday, dan adek ku Laily, yang selalu memberikan dukungan, perhatian, serta doa yang tulus dalam proses penyusunan skripsi ini.
3. Keluarga Besar Pondok Pesantren Tahfidzul Qur'an (PPTQ) Almustofa, tempat yang penuh berkah dan menjadi bagian penting dalam perjalanan hidup dan pembentukan kedisiplinan serta nilai-nilai keislaman penulis.

Semoga karya ini menjadi amal jariyah dan membawa manfaat bagi semua. Aamiin



ABSTRAK

Kedisiplinan adalah aspek fundamental dalam melakukan pembinaan karakter santri di pesantren. Selama ini pengawasan aktivitas harian masih dilakukan secara manual melalui pencatatan kehadiran dan pengamatan langsung, yang berisiko terhadap subjektivitas dan ketidakefisienan. Tujuan dari penelitian ini adalah untuk mendeteksi pola kedisiplinan santri berdasarkan data kehadiran dan aktivitas, dengan ruang lingkup pada santri aktif di lingkungan pesantren selama periode tiga tahun, yaitu dari tahun 2022 sampai dengan tahun 2024 Masehi yang secara ekuivalen berada pada rentang tahun 1444 sampai dengan 1446 Hijriyah. Penelitian ini menggunakan pendekatan data mining dengan algoritma K-Means Clustering. Penggunaan metode tersebut karena kemampuannya dalam mengelompokkan data secara otomatis berdasarkan kemiripan karakteristiknya, sehingga mampu mengidentifikasi klaster santri berdasarkan tingkat kedisiplinannya. Tahapan dari penelitian ini meliputi identifikasi masalah, studi literatur, pengumpulan data kehadiran dan aktivitas, pra-pemrosesan data, penerapan algoritma K-Means, evaluasi hasil klasterisasi, serta interpretasi dan analisis hasil. Hasil dari proses ini adalah menghasilkan beberapa kelompok santri dengan tingkat kedisiplinan berbeda, seperti santri disiplin tinggi dan kurang disiplin. Hasil dari pengelompokan ini diharapkan dapat memberikan gambaran yang objektif mengenai perilaku santri, sehingga dapat digunakan oleh pengasuh pesantren untuk merancang strategi pembinaan dan evaluasi secara lebih terarah. Penerapan algoritma K-Means pada penelitian ini menjadi solusi inovatif untuk mendukung manajemen pendidikan pesantren yang berbasis data. Pendekatan ini tidak hanya mempermudah dalam pengambilan keputusan, tetapi dapat memperkuat sistem pembinaan karakter di lingkungan pesantren.

Kata Kunci : kedisiplinan santri, kehadiran, aktivitas, data mining, K-Means Clustering

ABSTRACT

Discipline is a fundamental aspect of developing the character of students in Islamic boarding schools. Currently, monitoring of daily activities is still done manually through attendance recording and direct observation, which poses a risk of subjectivity and inefficiency. The purpose of this study is to detect patterns of student discipline based on attendance and activity data, focusing on active students in Islamic boarding schools over a three-year period, from 2022 to 2024 AD, which is equivalent to the years 1444 to 1446 Hijri. This study uses a data mining approach with the K-Means Clustering algorithm. This method is used because of its ability to automatically group data based on similar characteristics, thus identifying student clusters based on their level of discipline. The stages of this research include problem identification, literature review, attendance and activity data collection, data pre-processing, application of the K-Means algorithm, evaluation of clustering results, and interpretation and analysis of the results. The result of this process is the creation of several groups of students with varying levels of discipline, including highly disciplined and less disciplined students. The results of this grouping are expected to provide an objective picture of student behavior, which can then be used by Islamic boarding school administrators to design more targeted coaching and evaluation strategies. The application of the K-Means algorithm in this research provides an innovative solution to support data-driven Islamic boarding school education management. This approach not only facilitates decision-making but can also strengthen the character development system within the Islamic boarding school environment.

Keywords: student discipline, attendance, activities, data mining, K-Means Clustering

KATA PENGANTAR

Puji syukur kehadiran Allah Subhanahu wa Ta'ala atas segala limpahan rahmat, taufik, hidayah, dan inayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Klasterisasi Kedisiplinan Santri Berdasarkan Data Kehadiran dan Aktivitas Menggunakan Metode K-Means”** dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sains Data, Fakultas Sains dan Teknologi, Universitas Putra Bangsa Kebumen.

Tujuan umum dari penyusunan skripsi ini adalah untuk memberikan kontribusi ilmiah pada bidang analisis data, khususnya pada penerapan data mining di dalam konteks manajemen Pendidikan berbasis pesantren. Fokus utama pada penelitian ini adalah terletak pada pengolahan data kehadiran dan aktivitas santri untuk mendeteksi pola kedisiplinan, harapannya adalah dapat membantu pihak pesantren khususnya pengasuh dalam merancang strategi pembinaan santri yang lebih objektif dan efisien.

Isi dari skripsi ini adalah uraian tentang latar belakang pentingnya disiplin di dalam lingkungan pesantren, dalam proses analisis data metode yang digunakan menggunakan algoritma K-Means Clustering. Proses penelitian yang dilakukan melalui tahapan sistematis yang dijelaskan dalam bab-bab pada skripsi ini.

Kami mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyelesaian laporan ini terutama kepada:

1. Kedua Orangtua yang tercinta, yang selalu memberikan semangat, motivasi dan segala-galanya yang tak bisa diungkapkan dalam bentuk kata-kata.
2. Abah Rahmat, Umi Dewi, dan seluruh keluarga besar Pondok Pesantren Al-Mustofa Kebumen.

3. Saudara – saudara kandungku yang tercinta, Mas Huda, Mba Uni, Mas Baday, dan adekku Laily.
4. Bapak Dr. Gunarso Wiwoho, S.E., M.M. selaku Rektor Universitas Putra Bangsa Kebumen.
5. Bapak Sarjimin, S.Kom., M.Kom. selaku Kaprodi Sains Data.
6. Ibu Dian Rusvinasari, S.Kom., M.Kom. selaku dosen pembimbing.
7. Teman – teman satu Angkatan.
8. Almamater Universitas Putra Bangsa.
9. Dan semua pihak yang tidak bisa disebutkan satu persatu yang telah membantu.

Penelitian ini tidak akan terwujud tanpa bimbingan dan dukungan dari berbagai pihak. Kami berharap penelitian ini dapat memberikan wawasan yang berguna bagi pengasuh pesantren serta pihak-pihak terkait lainnya dalam upaya meningkatkan kedisiplinan santri di Pondok Pesantren Al-Mustofa Kebumen. Kritik dan saran yang membangun sangat kami harapkan untuk perbaikan di masa mendatang.

Akhir kata, semoga penelitian ini dapat memberikan manfaat dan kontribusi positif bagi pengembangan manajemen di Pondok Pesantren Al-Mustofa Kebumen.

Kebumen, 25 Juni 2025

Penulis,

DAFTAR ISI

HALAMAN JUDUL..... ii

HALAMAN PERSETUJUAN..... iii

HALAMAN PENGESAHAN UJIAN iv

HALAMAN BEBAS PLAGIARISME..... v

HALAMAN MOTTO vi

HALAMAN PERSEMBAHAN..... vii

ABSTRAK viii

ABSTRACT ix

KATA PENGANTAR x

DAFTAR ISI..... xii

DAFTAR TABEL..... xv

DAFTAR GAMBAR xvi

DAFTAR LAMPIRAN..... xvii

BAB I PENDAHULUAN 1

1.1. Latar Belakang Masalah 1

1.2. Rumusan Masalah 7

1.3. Batasan Masalah..... 8

1.4. Tujuan Penelitian..... 8

1.5. Manfaat Penelitian..... 9

BAB II KAJIAN PUSTAKA..... 10

2.1.	Tinjauan Teori	10
2.1.1.	<i>Grand Theory</i> :.....	10
2.1.2.	Teori Kedisiplinan.....	11
2.1.3.	Variabel X1: Kehadiran Santri.....	14
2.1.4.	Variabel X2: Aktivitas Santri.....	15
2.2.	Penelitian Terdahulu.....	16
2.3.	Hubungan Antar Variabel.....	19
2.4.	Model Empiris.....	21
BAB III METODE PENELITIAN.....		22
3.1.	Metode Penelitian.....	23
3.2.	Objek dan Subjek Penelitian	23
3.3.	Variabel Penelitian	25
3.4.	Definisi Operasional Variabel	26
3.5.	Instrumen atau Alat Pengumpulan Data.....	28
3.6.	Data dan Teknik Pengupulan Data	29
3.7.	Populasi dan Sampel	31
3.8.	Teknik Analisis.....	32
3.9.	Tahap Penelitian	33
BAB IV ANALISIS DAN PEMBAHASAN.....		36
4.1.	Deskripsi Data Penelitian	36
4.2.	Seleksi Data (Data Selection).....	48
4.3.	Pembersihan dan Persiapan Data (Data Cleaning & Preprocessing)	50
4.4.	Transformasi Data (Data Transformation)	58
4.5.	Proses Data Mining	60
4.6.	Implementasi Algoritma Data Mining.....	66
4.7.	Hasil Analisis dan Visualisasi	69

4.8. Pembahasan 77

BAB V KESIMPULAN DAN SARAN..... 80

5.1. Kesimpulan..... 80

5.2. Saran 81

DAFTAR PUSTAKA 82

LAMPIRAN..... 86



DAFTAR TABEL

Tabel II- 2 Penelitian Terdahulu	16
Tabel III- 1 Definisi Operasional untuk Semua Variabel	27
Tabel IV- 1 Sampel Data Identitas Santri	37
Tabel IV- 2 Sampel Data Kegiatan Madrasah	39
Tabel IV- 3 Sampel Data Ngaji Pagi.....	41
Tabel IV- 4 Sampel Data Ngaji Malam	43
Tabel IV- 5 Sampel Data Tahlilan	45
Tabel IV- 6 Sampel Data Manaqib	47
Tabel IV- 7 Tabel Pemberian Bobot.....	50
Tabel IV- 8 Sampel Pembobotan Kegiatan Madrasah.....	51
Tabel IV- 9 Sampel Pembobotan Ngaji Pagi	53
Tabel IV- 10 Sampel Pembobotan Ngaji Malam.....	54
Tabel IV- 11 Sampel Pembobotan Tahlilan.....	55
Tabel IV- 12 Sampel Pembobotan Manaqib.....	56

DAFTAR GAMBAR

Gambar III- 1 Flowchart Alur Metode Penelitian..... 22

Gambar IV- 1 Rumus untuk Menentukan Nilai Silhouette..... 61

Gambar IV- 2 Perhitungan Nilai Silhouette Scores pada Aplikasi Orange 62

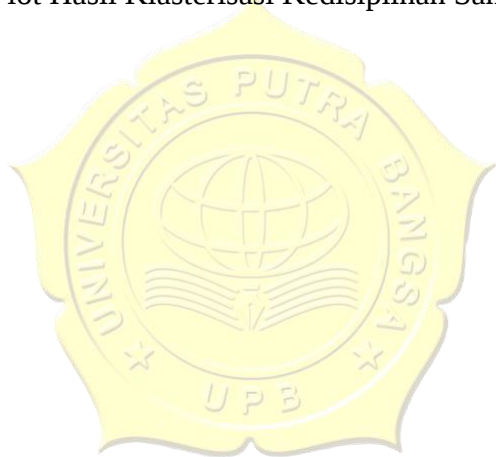
Gambar IV- 3 Rumus Euclidean Distance 64

Gambar IV- 4 Proses Implementasi Menggunakan Software Orange Data Mining.. 67

Gambar IV- 5 Hasil dari Cluster 1 70

Gambar IV- 6 Hasil dari Cluster 2 72

Gambar IV- 7 Scatter Plot Hasil Klasterisasi Kedisiplinan Santri..... 74



DAFTAR LAMPIRAN

Lampiran 1 Lokasi Pondok Pesantren Tahfidzul Qur'an Al-Mustofa Kuwayuhan,
Pejagoan, Kebumen..... 87

Lampiran 2 Sampel Identitas Santri Tahun..... 88

Lampiran 3 Sampel Data Kegiatan Madrasah 91

Lampiran 4 Sampel Data Ngaji Pagi..... 93

Lampiran 5 Sampel Data Ngaji Malam 95

Lampiran 6 Sampel Data Tahlilan 97

Lampiran 7 Sampel Data Manaqib 99

