BAB I
PENDAHULUAN

1.1 Latar Belakang

Pondok pesantren memiliki peran penting dalam pembentukan karakter santri di Indonesia. Sebagai lembaga pendidikan berbasis Islam, pesantren tidak hanya menyediakan pendidikan akademik tetapi juga pendidikan karakter yang mencakup nilai-nilai moral, disiplin, dan keagamaan. Lingkungan pesantren menjadi ruang yang unik di mana tradisi, nilai spiritual, dan kedisiplinan diterapkan secara konsisten. Dengan berlandaskan prinsip-prinsip Islam, pesantren menjadi tempat yang membentuk santri untuk menjadi individu yang tidak hanya cerdas secara intelektual tetapi juga matang secara emosional dan spiritual. Di tengah tantangan modernisasi dan globalisasi, peran pesantren menjadi semakin krusial dalam mempertahankan nilai-nilai luhur budaya dan agama, sekaligus menyiapkan generasi muda yang kompetitif dan berkarakter.

Berdasarkan data Kementerian Agama Republik Indonesia tahun 2023, jumlah pondok pesantren di Indonesia telah mencapai lebih dari 36.000 lembaga dengan total santri aktif melebihi 4,2 juta orang. Angka ini menjadikan pesantren sebagai salah satu sistem pendidikan terbesar dan paling tersebar di seluruh nusantara. Meskipun demikian, pengelolaan data santri di sebagian besar lembaga pesantren masih dilakukan secara konvensional dan belum memanfaatkan pendekatan berbasis data (data-driven) secara optimal. Kondisi ini menjadi tantangan tersendiri, terutama dalam merancang strategi pembinaan yang tepat dan terukur bagi ribuan santri dengan karakteristik yang beragam.

SEKOLAH PASCASARJANA

Di era transformasi digital, ketersediaan data yang besar membuka peluang baru untuk memahami karakteristik peserta didik secara lebih mendalam. Dalam konteks pesantren, data perilaku, penilaian santri, dan rekam jejak keagamaan santri sesungguhnya merupakan sumber informasi berharga yang dapat diolah untuk mendukung pengambilan keputusan berbasis bukti *(evidence-based decision making)*. Namun, hingga saat ini penelitian yang secara khusus mengintegrasikan pendekatan machine learning ke dalam analisis profil santri di

lingkungan pesantren masih sangat terbatas, sehingga potensi besar ini belum banyak dimanfaatkan.

Penelitian-penelitian sebelumnya yang menggunakan teknik klasterisasi dalam konteks pendidikan umumnya berfokus pada mahasiswa di perguruan tinggi, dengan variabel yang terbatas pada aspek akademik semata. Belum banyak penelitian yang mengintegrasikan dimensi karakter, gaya belajar, dan indikator keagamaan secara serentak, terlebih dalam konteks lembaga pendidikan Islam seperti pesantren yang memiliki karakteristik unik. Celah penelitian (research gap) inilah yang menjadi salah satu motivasi utama penelitian ini, yakni untuk mengembangkan model profiling santri yang bersifat multidimensi dengan memanfaatkan pendekatan hybrid clustering.

Namun, proses pembentukan karakter tidak hanya ditentukan oleh lingkungan pesantren itu sendiri. Faktor individual, seperti gaya belajar, juga memberikan kontribusi signifikan dalam proses pendidikan. Gaya belajar santri yang bervariasi meliputi visual, auditorial, dan kinestetik menjadi tantangan tersendiri bagi pesantren untuk menciptakan pendekatan pendidikan yang inklusif. Dalam praktiknya, gaya belajar yang sesuai dapat meningkatkan efektivitas proses pembelajaran dan membantu santri untuk menyerap nilai-nilai karakter yang diajarkan. Sebaliknya, jika gaya belajar tidak selaras dengan pendekatan pembelajaran, potensi santri dalam menginternalisasi pendidikan karakter dapat terhambat. Oleh karena itu, memahami interaksi antara gaya belajar dan lingkungan pesantren menjadi penting dalam merancang strategi pendidikan yang efektif.

Untuk menjawab tantangan tersebut, penelitian ini menggunakan pendekatan analisis data berbasis klastering untuk mengeksplorasi hubungan antara lingkungan pesantren, gaya belajar, dan karakter santri. Teknik seperti *Hierarchical Clustering* dan K-Means dipilih karena kemampuan keduanya untuk mengidentifikasi pola dalam data yang kompleks. *Hierarchical Clustering* digunakan karena metode ini tidak memerlukan penentuan jumlah klaster di awal, sehingga memudahkan eksplorasi hubungan antar data dengan struktur hierarkis. Dendrogram yang dihasilkan memungkinkan visualisasi hubungan antar karakteristik santri sebelum menentukan klaster definitif. Setelah jumlah klaster

SEKOLAH PASCASARJANA

optimal diperoleh menggunakan Silhouette Score, K-Means *Clustering* diterapkan untuk membentuk klaster yang lebih terdefinisi. Pemilihan metode ini didasarkan pada efisiensi dan kecocokan dengan karakteristik data yang digunakan dalam penelitian.

Metode lain seperti DBSCAN dan Gaussian Mixture Model (GMM) juga dipertimbangkan tetapi tidak dipilih karena beberapa alasan. DBSCAN lebih cocok untuk data yang memiliki distribusi tidak beraturan dengan banyak noise, sedangkan dalam penelitian ini, karakter santri diperkirakan memiliki pola struktur yang lebih jelas. GMM, yang berbasis pendekatan probabilistik, kurang sesuai karena penelitian ini tidak mengasumsikan bahwa karakter santri mengikuti distribusi normal tertentu. Dengan demikian, kombinasi *Hierarchical Clustering* dan K-Means merupakan pendekatan yang optimal dalam memahami hubungan antara lingkungan pesantren, gaya belajar, dan pembentukan karakter santri.

Pendekatan berbasis klaster ini tidak hanya relevan dalam pendidikan, tetapi juga telah digunakan dalam berbagai disiplin ilmu untuk menangani data yang kompleks. Misalnya, (Govender & Sivakumar, 2020) menunjukkan bahwa metode klastering mampu mengidentifikasi pola-pola tersembunyi dalam data yang kompleks, memberikan wawasan yang lebih mendalam tentang struktur data yang mendasarinya, khususnya dalam konteks pendidikan. Mereka membuktikan bahwa teknik ini efektif untuk memahami hubungan dalam data yang tidak langsung terlihat, seperti dalam penelitian pendidikan yang melibatkan banyak variabel.

Selain itu, penelitian oleh Niu dkk. (2021) mengaplikasikan kombinasi *Hierarchical Clustering* dan K-Means untuk menganalisis pola energi pada sistem suplai energi di pulau-pulau. Pendekatan ini berhasil mengidentifikasi klasifikasi yang relevan, menunjukkan bagaimana metode ini dapat diterapkan untuk memahami data kompleks di berbagai bidang, termasuk pendidikan. Penelitian ini menegaskan bahwa teknik klastering tidak hanya efektif dalam mengungkap pola tersembunyi tetapi juga memberikan solusi berbasis data untuk permasalahan praktis.

Dengan demikian, penelitian ini mengadopsi pendekatan multidimensi untuk menganalisis pengaruh lingkungan pesantren dan gaya belajar terhadap profil

SEKOLAH PASCASARJANA

karakter santri. Penggunaan analisis klaster seperti *Hierarchical Clustering* dan K-Means memberikan peluang untuk menggali hubungan kompleks antara faktor-faktor tersebut dan mengidentifikasi klaster-karakter yang berbeda, yang belum banyak dikaji secara menyeluruh dalam penelitian sebelumnya.

Hasil dari penelitian ini diharapkan dapat memberikan wawasan baru dalam memahami dinamika pembentukan karakter santri dan memberikan kontribusi praktis dalam pengelolaan pendidikan di pesantren. Temuan penelitian ini dapat digunakan untuk mengembangkan kurikulum yang responsif terhadap kebutuhan santri, menciptakan strategi pendidikan berbasis data yang efektif, dan melatih pendidik dalam memahami serta menyesuaikan pendekatan pembelajaran dengan gaya belajar santri. Dengan demikian, penelitian ini tidak hanya memperkaya literatur ilmiah tetapi juga menawarkan solusi praktis yang relevan bagi pengelola pesantren, pendidik, dan pembuat kebijakan pendidikan di Indonesia.

1.2 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengembangkan model *profiling* karakter santri berbasis *machine learning* dengan pendekatan klasterisasi *hybrid* yang mengintegrasikan *Hierarchical Clustering* dan *K-Means*.
2. Mengelompokkan santri berdasarkan tiga kelompok atribut, yaitu dimensi karakter individual (kedisiplinan, tanggung jawab, kejujuran, empati, dan interaksi sosial), preferensi gaya belajar (visual, auditori, kinestetik), serta indikator performa akademik dan keagamaan.
3. Menentukan jumlah klaster berdasarkan struktur alami data yang teridentifikasi melalui analisis dendrogram pada *Hierarchical Clustering*, sebagai dasar penentuan jumlah klaster yang digunakan dalam tahap *K-Means*.
4. Menghasilkan profil karakter santri yang representatif dan bermakna secara kontekstual, sebagai dasar perancangan strategi pembinaan yang lebih adaptif, personal, dan berbasis data.

SEKOLAH PASCASARJANA

1.3 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi sebagai berikut :

1. Peningkatan Pemahaman Karakter Santri :

   Dengan mengintegrasikan analisis lingkungan pesantren dan gaya belajar menggunakan metode *Hierarchical Clustering* dan *k-means*, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai pola karakteristik santri, sehingga mendukung proses pendidikan yang lebih efektif.

2. Pengembangan Metode Analisis Karakter :

   Penelitian ini dapat memberikan alternatif metodologi baru dalam pengelompokkan data karakteristik santri berbasis klaster. Integrasi metode *Hierarchical Clustering* dan *k-means* diharapkan memberikan hasil yang lebih akurat dan relevan untuk mendukung evaluasi dan pengembangan sistem pendidikan pesantren.

SEKOLAH PASCASARJANA