



Analisis Klasifikasi Kesiapan Digital Desa Menggunakan Decision Tree dan Pemetaan Spasial

Hafidlotul Fatimah Ahmad^{1,*}, Aulia Rizki Firdawanti¹, Nur Agustiani³

¹ Sekolah Sains Data, Matematika, dan Informatika, Prodi Ilmu Komputer, Institut Pertanian Bogor, Bogor, Indonesia

² Sekolah Sains Data, Matematika, dan Informatika, Prodi Statistika dan Sains Data, Institut Pertanian Bogor, Bogor, Indonesia

³ Sekolah Sains Data, Matematika, dan Informatika, Program Studi Aktuaria, Institut Pertanian Bogor, Bogor, Indonesia

Email: ^{1,*}hafidlotulftm@apps.ipb.ac.id, ²auliafirda@apps.ipb.ac.id, ³nur_agustiani@apps.ipb.ac.id

Email Penulis Korespondensi: hafidlotulftm@apps.ipb.ac.id

Abstrak—Transformasi digital di tingkat desa merupakan elemen strategis dalam upaya pemerataan pembangunan dan peningkatan kualitas pelayanan publik. Namun, tingkat kesiapan digital antarwilayah belum merata secara signifikan. Penelitian ini bertujuan mengklasifikasikan tingkat kesiapan digital (digital readiness) desa-desa pada kabupaten/kota di Provinsi Jawa Barat dengan memanfaatkan data terbuka pada Open Data Jabar (opendata.jabarprov.go.id) terkait jumlah desa digital, akses internet, dan kategori strata desa. Algoritma klasifikasi Decision Tree digunakan untuk memetakan wilayah ke dalam dua kelas readiness: ‘tinggi’ dan ‘rendah’. Hasil pemodelan menunjukkan bahwa jumlah desa digital dan persentase desa dengan akses internet merupakan variabel yang paling berpengaruh dalam klasifikasi. Meskipun infrastruktur internet tersedia di sebagian besar wilayah, hal tersebut tidak selalu berbanding lurus dengan status digitalisasi desa. Kabupaten dengan akses internet tinggi namun jumlah desa mandiri rendah tetap tergolong readiness rendah. Akurasi model mencapai 83%, namun performa pada kelas readiness tinggi masih rendah karena distribusi data yang tidak seimbang. Visualisasi spasial turut digunakan untuk menegaskan ketimpangan kesiapan digital antarwilayah yang diamati pada penelitian ini. Penelitian ini memberikan kontribusi awal dalam pemetaan digital readiness desa berbasis pendekatan machine learning di Indonesia.

Kata Kunci: Decision Tree; Digital Readiness; Desa Digital; Machine Learning; Pemetaan Spasial

Abstract—Digital transformation at the village level is a strategic element in promoting equitable development and improving public service delivery. However, the level of digital readiness across regions remains uneven. This study aims to classify the digital readiness of villages in West Java Province by utilizing data from Open Data Jabar (opendata.jabarprov.go.id) related to the number of digital villages, internet access, and village development strata. A Decision Tree classification algorithm was employed to categorize regions into two readiness classes: high and low. The modeling results indicate that the number of self-reliant (mandiri) villages and the percentage of villages with internet access are the most influential variables in the classification. Although internet infrastructure is available in most areas, it does not always correspond to the level of village digitalization. Districts with high internet access but a low number of self-reliant villages are still classified as having low readiness. The model achieved an accuracy of 83%, although its performance in identifying the high readiness class was limited due to class imbalance in the dataset. Spatial visualization was also used to highlight regional disparities in digital readiness. This study provides an early contribution to digital readiness mapping of villages using a machine learning approach in Indonesia.

Keywords: Decision Tree; Digital Readiness; Digital Village; Machine Learning; Spatial Mapping

1. PENDAHULUAN

Transformasi digital telah menjadi pilar penting dalam pembangunan di berbagai sektor, termasuk pada tingkat pemerintahan desa. Program Desa Digital yang dicanangkan oleh Kementerian Komunikasi dan Digital (Komdigi) bertujuan untuk mendorong pemanfaatan teknologi informasi dalam layanan publik dan pemberdayaan masyarakat desa [1]. Meski demikian, tantangan pemerataan masih terlihat jelas. Survei Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) tahun 2024 menunjukkan bahwa 17,4% masyarakat desa belum terkoneksi internet [2], sementara survei Kementerian Kominfo (2021) mencatat 50,82% responden di kawasan pedesaan memiliki tingkat literasi digital di bawah rata-rata nasional [3]. Kondisi ini memunculkan pertanyaan mengenai kesiapan setiap daerah dalam mengimplementasikan program secara efektif. Kesiapan digital desa atau *digital readiness* mengacu pada kondisi suatu wilayah desa dalam hal infrastruktur, kapasitas manusia, dan keberadaan layanan digital yang memungkinkan proses transformasi digital berjalan optimal. Menurut Chinoracky *et al.*, (2025) [4] *digital readiness* memerlukan infrastruktur fisik dan digital seperti jaringan, pusat data, listrik, dan perangkat pendukung. Selaras dengan itu, Duanmu (2025) menggarisbawahi bahwa infrastruktur merupakan pondasi penting dalam mengurangi ketimpangan digital [5].

Keberhasilan transformasi digital tidak hanya ditentukan oleh ketersediaan teknologi, tetapi juga oleh kesiapan wilayah dalam mengadopsinya secara menyeluruh, termasuk kesiapan infrastruktur dan sumber daya manusia [6]. Konsep ini menjadi semakin relevan ketika diterapkan pada konteks negara berkembang seperti Indonesia, di mana kesenjangan digital antar wilayah masih cukup tinggi [7], [8]. Di tingkat nasional, Pemerintah Indonesia melalui Kementerian Komunikasi dan Digital telah meluncurkan program Desa Digital sebagai upaya untuk mendorong akselerasi pembangunan berbasis teknologi informasi di pedesaan. Namun, implementasi program ini menghadapi tantangan besar, terutama karena belum meratanya infrastruktur digital [9] dan kapasitas lokal yang mendukung. Berdasarkan Indeks Masyarakat Digital Indonesia (IMDI) 2023, skor infrastruktur desa hanya mencapai 57,09, sedangkan skor pemberdayaan dan ketenagakerjaan desa jauh lebih rendah, masing-masing 26,19 dan 32,14, yang menunjukkan rendahnya pemanfaatan teknologi digital di pedesaan [10].

Salah satu provinsi yang menjadi sorotan adalah Jawa Barat, yaitu provinsi dengan jumlah desa terbanyak ke-5 di Indonesia [11]. Provinsi ini memiliki keragaman karakteristik geografis, sosial, dan ekonomi yang tinggi, sehingga



merepresentasikan beragam kondisi yang mempengaruhi kesiapan digital di tingkat desa. Implementasi komunikasi digital dalam pelayanan publik di Jawa Barat menunjukkan adanya dinamika yang menarik, dengan adanya potensi pengembangan sekaligus ruang perbaikan pada tingkat administratif [12]. Selain itu, keterlibatan banyak kota/kabupaten di Jawa Barat dalam Program 100 Smart Cities [13] menjadikan provinsi ini menjadi lokasi yang relevan untuk melakukan kajian terkait kesiapan digital. Data dari Open Data Jabar (opendata.jabarprov.go.id) menunjukkan jumlah desa digital di Jawa Barat sangat bervariasi antar kabupaten/kota, begitu juga dengan tingkat akses internet dan jumlah desa mandiri. Ketimpangan ini dapat menjadi penghambat utama dalam pemerataan digitalisasi desa, terutama dalam konteks pelayanan publik berbasis digital. Oleh karena itu, penting untuk melakukan pemetaan tingkat kesiapan digital desa secara sistematis guna mendukung perumusan kebijakan berbasis data. Pemetaan *digital readiness* dapat membantu pemerintah menghindari kebijakan yang bersifat generik dan tidak tepat sasaran, sekaligus mengoptimalkan alokasi sumber daya ke wilayah yang paling membutuhkan [14]. Selain itu, pemetaan yang komprehensif berfungsi sebagai *baseline* data dalam proses monitoring dan evaluasi [15] transformasi digital desa. Dengan tersedianya data *digital readiness* yang terbuka dan akurat, proses pengambilan keputusan dapat berlangsung secara lebih transparan, efisien, dan partisipatif, sesuai dengan prinsip tata kelola pemerintahan berbasis digital.

Beberapa penelitian sebelumnya telah membahas topik terkait kesiapan digital desa di Indonesia, namun pendekatannya umumnya masih bersifat deskriptif dan belum memanfaatkan metode *machine learning*. Misalnya, penelitian oleh Musfikar *et al.*, (2022) [16] mengevaluasi kesiapan desa menuju *Smart Village* di Kabupaten Aceh Besar, menggunakan metode deskriptif dan indikator sosial-budaya. Sementara itu, Aditianto *et al.*, (2024) [17] meneliti mengenai tingkat kesiapan pemanfaatan Sistem Informasi Desa (SID) di Kabupaten Purworejo menggunakan metode SmartPLS. Penelitian ini menunjukkan bahwa kesiapan desa dalam memanfaatkan SID masih rendah. Mayasari *et al.*, (2023) [18] mengevaluasi kesiapan pemerintah desa dalam mengadopsi e-Government menggunakan kerangka STOPE (*Strategy, Technology, Organization, People, Environment*). Hasilnya menunjukkan bahwa tiga dari lima domain (strategi, organisasi, dan lingkungan) dinyatakan sangat siap, sementara aspek teknologi dan SDM masih perlu penguatan. Penelitian di Desa Semaya dan Pegiringan, Kabupaten Pemalang oleh Subhan dan Pratama (2022) [19] juga melakukan analisis terkait kesiapan desa dalam penerapan E-Government berdasarkan aspek sistem data, regulasi, kelembagaan, sumber daya manusia, dan infrastruktur TIK. Namun, seluruh penelitian tersebut belum mengadopsi pendekatan berbasis *machine learning* dalam proses analisisnya, sehingga masih terbatas pada metode deskriptif atau model statistik konvensional. Penerapan *machine learning* dalam konteks pembangunan desa digunakan pada penelitian Sitorus *et al.*, (2022) [20], yang menggunakan metode klasifikasi untuk memetakan daerah rawan banjir di Desa Pahlawan, Kabupaten Batu Bara. Namun, fokus penelitian tersebut lebih kepada mitigasi bencana dan belum menyoroti isu kesiapan digital desa secara langsung. Dengan demikian, penelitian ini menjadi kontribusi awal dalam menerapkan pendekatan berbasis *machine learning* untuk pemetaan *digital readiness* di Indonesia.

Penggunaan pendekatan *machine learning* memungkinkan dilakukannya klasifikasi secara sistematis dan berbasis data, yang tidak hanya meningkatkan objektivitas analisis, tetapi juga mampu mengungkap pola-pola tersembunyi yang sulit diidentifikasi melalui pendekatan konvensional. Dalam penelitian ini, algoritma *Decision Tree* dipilih karena memiliki tingkat interpretabilitas yang tinggi dan mudah divisualisasikan [21], sehingga cocok digunakan dalam konteks kebijakan publik. Selain itu, algoritma *Decision Tree* cukup fleksibel dalam menangani data dengan skala kecil [22], serta mampu menangkap interaksi non-linier antar variabel yang mungkin terlewat oleh pendekatan konvensional. Oleh karena itu, penelitian ini bertujuan untuk membangun model klasifikasi yang dapat memetakan *digital readiness* desa di Jawa Barat menggunakan metode *Decision Tree*. Dengan klasifikasi ini, diharapkan dapat diperoleh gambaran wilayah mana saja yang perlu mendapatkan perhatian dan prioritas dalam program digitalisasi desa.

2. METODOLOGI PENELITIAN

2.1 Digital Readiness

Digital readiness merupakan kondisi kesiapan multi-dimensi yang meliputi infrastruktur/teknologi, keterampilan & literasi digital, proses/organisasi, dan dukungan kelembagaan. Kesiapan ini menjadi faktor penentu keberhasilan transformasi layanan publik maupun pembangunan wilayah secara berkelanjutan. Secara konseptual, *digital readiness* dipandang sebagai kombinasi kemampuan *sensemaking* teknologi, kelincahan (*agility*), dan implementasi teknologi baru sehingga perlu mempertimbangkan dimensi manusia, proses, dan teknologi [23]. Kerangka dan tinjauan mutakhir menunjukkan bahwa penilaian *readiness* kerap beririsan dengan model kematangan (*maturity*) untuk memetakan posisi awal dan menyusun peta jalan perbaikan organisasi/pemerintah, menekankan pentingnya pondasi infrastruktur TIK, *data center*, listrik andal, serta tata kelola data/keamanan [24]. Dalam konteks pedesaan, *digital readiness* memiliki karakteristik khusus berupa akses/infrastruktur, literasi, dan layanan publik lokal, yang secara bersama menentukan kesenjangan antarwilayah serta prioritas intervensi kebijakan untuk memajukan *digital readiness* desa [25].

2.2 Desa Digital

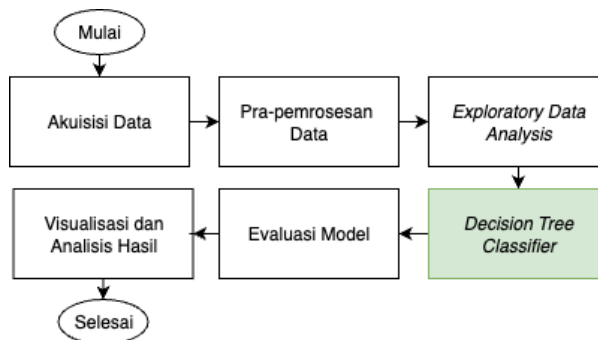
Konsep desa digital (*smart village*) merupakan implementasi nyata dari prinsip *digital readiness* di tingkat pedesaan. Desa digital atau *smart village* didefinisikan sebagai desa yang mampu secara inovatif memanfaatkan teknologi informasi dan komunikasi (TIK) untuk meningkatkan kualitas hidup, efisiensi layanan, serta daya saing ekonomi, sosial, dan lingkungan sembari tetap memelihara kearifan lokal. Contohnya dapat dilihat pada studi Kasus Kalurahan Panggungharjo,



Yogyakarta, yang menyorot pemanfaatan Sistem Informasi Desa (SID) untuk memperkuat interaksi antara pemerintah desa, masyarakat, dan lingkungan sekitar [26]. Lebih lanjut, kerangka konseptual desa digital Indonesia menekankan tiga dimensi utama (*smart governance*, *smart community*, dan *smart environment*) sebagai landasan inovasi desa digital yang berakar pada kelembagaan lokal, struktur sosial desa, dan konteks lingkungan pedesaan [27].

2.3 Tahapan Penelitian

Tahapan pada penelitian ini dimulai dari pengumpulan data yang relevan, dilanjutkan dengan pra-pemrosesan untuk menjamin kualitas data, serta analisis eksploratori guna memahami karakteristik dan pola awal yang terkandung di dalamnya. Tahapan berikutnya mencakup pembangunan model klasifikasi menggunakan algoritma *Decision Tree*, yang kemudian dievaluasi performanya sebelum hasilnya divisualisasikan dan dianalisis secara menyeluruh. Rangkaian tahapan penelitian tersebut disajikan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.3.1 Akuisisi Data

Pada tahap akuisisi data, dilakukan proses pengumpulan data dari portal Open Data Jabar dengan fokus pencarian pada kata kunci “desa digital” dan “desa”. Data yang diperoleh mencakup berbagai indikator terkait digitalisasi desa, meliputi jumlah desa digital, tingkat akses internet, serta klasifikasi wilayah desa di seluruh kabupaten/kota di Jawa Barat.

2.3.2 Pra-pemrosesan Data

Data yang telah dikumpulkan kemudian dipra-proses untuk menyiapkan data menjadi bentuk yang layak digunakan dalam proses analisis dan pemodelan. Langkah pertama yang dilakukan adalah menggabungkan beberapa *dataset* yang diperoleh dari tahap akuisisi. Selanjutnya dilakukan transformasi terhadap variabel kategorik menjadi format biner menggunakan metode *one-hot encoding* untuk memastikan model interpretatif [28]. Selain itu, dibuat dua variabel turunan, yaitu persentase desa digital dan persentase desa dengan akses internet, yang dihitung berdasarkan proporsi terhadap jumlah total desa di masing-masing kabupaten/kota. Tahap akhir pra-pemrosesan mencakup pembersihan data dengan cara menghapus baris yang memiliki nilai kosong secara dominan pada sebagian fitur, sedangkan untuk kasus nilai kosong yang terbatas pada beberapa fitur tertentu, akan dilakukan imputasi guna mempertahankan integritas dan kelengkapan data.

2.3.3 Exploratory Data Analysis

Tahap eksplorasi data bertujuan untuk memperoleh pemahaman awal terhadap distribusi data serta potensi hubungan antar variabel. Proses ini dilakukan melalui penerapan teknik visualisasi yang disesuaikan dengan karakteristik masing-masing fitur dalam dataset. Selain itu, dilakukan juga analisis korelasi untuk mengidentifikasi kontribusi dan keterkaitan antar variabel.

2.3.4 Decision Tree Classifier

Sebelum dilakukan pemodelan klasifikasi, akan dilakukan penetapan skema kategorisasi tingkat kesiapan digital kabupaten/kota berdasarkan variabel persentase desa digital. Akan dilakukan penentuan ambang batas (*threshold*) untuk membuat dua kategori: 1) *readiness* tinggi, 2) *readiness* rendah. Penetapan ambang batas ini bersifat eksploratif untuk memperoleh label target yang akan digunakan sebagai variabel dependen dalam proses pelatihan model klasifikasi. Pemodelan klasifikasi dilakukan menggunakan algoritma *Decision Tree* karena memiliki tingkat interpretabilitas yang tinggi dan kemampuan visualisasi yang intuitif dalam bentuk pohon keputusan [29]. Dataset dibagi menjadi dua subset, yaitu data pelatihan dan data pengujian dengan proporsi 70:30 [30] untuk mengevaluasi kinerja model secara objektif.

2.3.5 Evaluasi Model

Untuk mengukur kinerja model klasifikasi, digunakan beberapa metrik evaluasi yaitu akurasi, *precision*, *recall* dan *F1-score*. Akurasi digunakan untuk melihat proporsi prediksi yang benar secara keseluruhan. *Precision* dan *recall* digunakan untuk melihat distribusi kelas yang tidak seimbang. *F1-score* yang merupakan penggabungan antara *presisi* dan *recall* [31] digunakan untuk menilai performa model secara menyeluruh terhadap setiap kelas.



2.3.6 Visualisasi dan Analisis Hasil

Hasil klasifikasi divisualisasikan menggunakan pohon keputusan hasil *Decision Tree* dan representasi spasial dalam bentuk peta wilayah provinsi Jawa Barat. Visualisasi pohon keputusan digunakan untuk menggambarkan secara hierarkis aturan-aturan klasifikasi yang dihasilkan oleh model. Sementara itu, representasi spasial digunakan untuk memetakan hasil visualisasi ke dalam bentuk peta provinsi Jawa Barat, sehingga distribusi kelas dapat diamati secara geografis.

3. HASIL DAN PEMBAHASAN

3.1 Data

Penelitian ini menggunakan data sekunder yang diperoleh dari portal resmi **Open Data Jabar** tahun 2021. Terdapat 3 dataset yang digunakan seperti yang dapat dilihat pada Tabel 1.

Tabel 1. Dataset yang digunakan

Nomor	Judul Dataset	Fitur
1	Jumlah Kawasan Desa Digital Berdasarkan Kabupaten/Kota di Jawa Barat (https://opendata.jabarprov.go.id/id/dataset/jumlah-kawasan-des-digital-berdasarkan-kabupatenkota-di-jawa-barat)	jumlah_desa, satuan, tahun
2	Jumlah Desa yang Memiliki Akses Internet Berdasarkan Status Keberadaan di Jawa Barat (https://opendata.jabarprov.go.id/id/dataset/jumlah-des-digital-berdasarkan-status-keberadaan-di-jawa-barat)	status_keberadaan, jumlah_desa, satuan, tahun
3	Jumlah Desa Berdasarkan Status Strata Desa di Jawa Barat (https://opendata.jabarprov.go.id/id/dataset/jumlah-des-digital-berdasarkan-status-strata-des-digital-di-jawa-barat)	strata_desa, jumlah_desa, satuan, tahun

3.2 Exploratory Data Analysis

Statistik deskriptif untuk variabel jumlah desa digital dan kesediaan akses internet dapat dilihat pada Tabel 2. Rata-rata jumlah desa digital per kabupaten/kota adalah sekitar 7.42 desa, dengan standar deviasi sebesar 7.75 menunjukkan adanya variasi cukup besar antar wilayah. Nilai minimum 0 menunjukkan ada kabupaten/kota yang belum memiliki desa digital sama sekali. Nilai Q1 dan Q2 sebesar 2 dan 3 desa mengindikasikan bahwa sebagian besar wilayah memiliki jumlah desa digital yang relatif rendah. Nilai maksimum sebesar 27 desa (dibandingkan dengan *mean* 7.42) menegaskan adanya kesenjangan antar wilayah dalam hal digitalisasi desa, baik dari sisi program, anggaran, maupun kesiapan sumber daya manusia di tingkat lokal, yang bisa disebabkan oleh faktor geografis, kebijakan daerah, atau tingkat adopsi teknologi yang belum merata.

Pada variabel kesediaan akses internet, rata-rata desa dengan internet berjumlah 229.11, sedangkan rata-rata desa tanpa internet berjumlah 50.47. Hal ini menunjukkan mayoritas desa di provinsi Jawa Barat yang memiliki akses internet. *Median* (Q2) jumlah desa dengan akses internet adalah 253 desa, dengan Q3 hampir mencapai 297 desa, menunjukkan bahwa separuh wilayah memiliki infrastruktur internet yang cukup baik. Namun demikian, masih terdapat kabupaten/kota dengan 125 desa yang tidak memiliki akses internet, mencerminkan masih adanya tantangan dalam pemerataan akses digital. Selain itu, nilai standar deviasi pada kedua variabel internet (dengan dan tanpa akses) masing-masing sebesar 91.63 dan 32.41, mengindikasikan adanya variasi cukup besar antar wilayah. Hal ini penting dicermati karena menunjukkan adanya ketidakseimbangan sebaran infrastruktur digital, yang dapat menjadi pertimbangan dalam tahap analisis lebih lanjut maupun perumusan kebijakan pemerataan akses digital di tingkat desa.

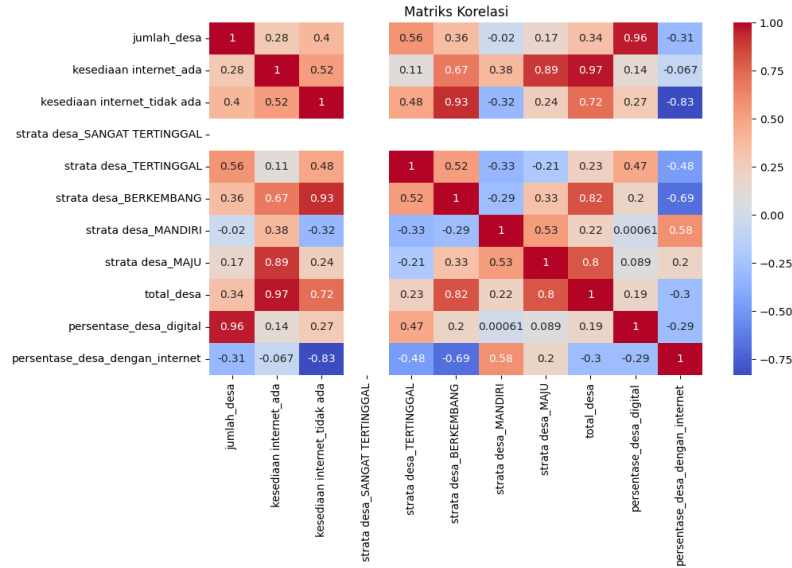
Tabel 2. Statistik deskriptif variabel jumlah desa digital dan akses internet

No	Statistik	Jumlah desa digital	Desa dengan internet	Desa tanpa internet
1	<i>Count</i>	19.00	19.00	19.00
2	<i>Mean</i>	7.42	229.11	50.47
3	<i>Std Dev</i>	7.75	91.63	32.41
4	<i>Minimum</i>	0.00	15.00	1.00
5	25% (Q1)	2.00	174.00	28.00
6	<i>Median</i> (Q2)	3.00	253.00	40.00
7	75% (Q3)	14.00	296.50	67.50
8	<i>Maximum</i>	27.00	368.00	125.00

Pada Gambar 2 dapat dilihat matriks korelasi antar variabel. Variabel *persentase_desa_digital* memiliki korelasi tinggi dengan variabel *jumlah_desa* yaitu sebesar 0.96, hal ini menandakan semakin banyak jumlah desa di kabupaten/kota maka semakin tinggi pula persentase desa digital yang tercatat. Korelasi negatif antara variabel



persentase_desa_dengan_internet dan strata desa BERKEMBANG (-0.69) serta strata desa TERTINGGAL (-0.48) menunjukkan desa dengan status tertinggal dan berkembang cenderung memiliki akses internet yang rendah. Sebaliknya, terdapat korelasi positif antara variabel persentase_desa_dengan_internet dengan variabel strata desa MANDIRI (0.58) menunjukkan bahwa desa dengan status pembangunan yang lebih tinggi cenderung memiliki infrastruktur digital yang lebih baik.

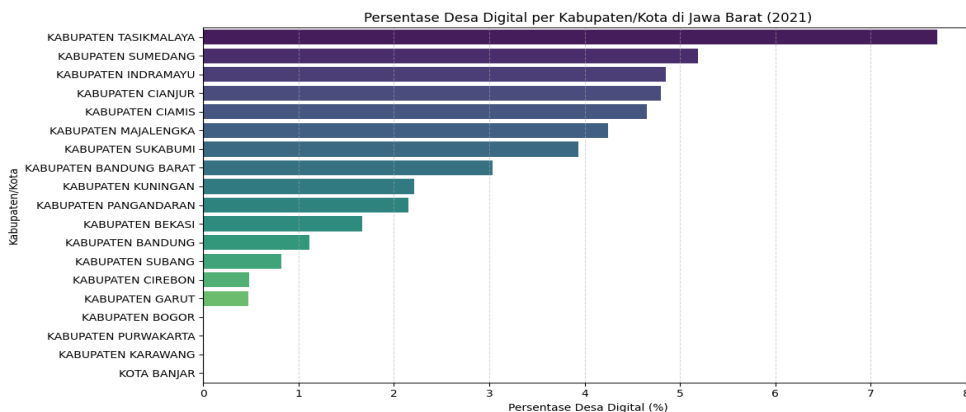


Gambar 2. Heatmap korelasi antar variabel

Lebih lanjut, korelasi tinggi antara variabel jumlah_desa dan variabel persentase_desa_digital (0.96) juga membuka peluang bahwa kabupaten yang memiliki perhatian lebih terhadap digitalisasi cenderung memiliki basis administratif (jumlah desa) yang besar. Hal ini dapat disebabkan oleh adanya dukungan kelembagaan dan sumber daya manusia yang lebih luas, sehingga lebih mampu mengadopsi teknologi digital secara merata dan berkelanjutan di berbagai sektor pemerintahan dan layanan masyarakat desa.

Dari total 26 kabupaten/kota di Provinsi Jawa Barat, setelah melalui tahap pra-pemrosesan data, hanya terdapat 19 wilayah yang dapat dianalisis lebih lanjut. Sebagian besar diantaranya merupakan kabupaten, sementara data untuk wilayah kota tidak tersedia atau tidak lengkap sehingga tidak disertakan dalam analisis. Distribusi persentase desa digital per kabupaten/kota di provinsi Jawa Barat dapat dilihat pada Gambar 3. Kabupaten Tasikmalaya tercatat memiliki persentase tertinggi yaitu di atas 7%, diikuti oleh Kabupaten Sumedang dengan persentase di atas 5%. Walaupun persentase desa digital memiliki korelasi yang positif dengan jumlah desa digital, perlu dicermati bahwa rendahnya persentase desa digital tidak selalu berbanding lurus dengan jumlah desa digital secara absolut. Sebagai contoh, Kabupaten Bandung memiliki jumlah desa yang banyak, namun persentase digitalisasinya tetap rendah, yang berarti hanya sebagian kecil desa yang telah tersentuh program tersebut.

Fenomena ini menyoroti pentingnya mempertimbangkan proporsi dalam analisis, bukan hanya angka absolut. Pendekatan berbasis persentase memberikan perspektif yang lebih adil dalam membandingkan wilayah dengan jumlah desa yang sangat berbeda. Grafik ini memperkuat pentingnya klasifikasi digital readiness, karena tidak semua kabupaten/kota memulai dari titik yang sama. Intervensi kebijakan dan alokasi sumber daya perlu mempertimbangkan fakta bahwa distribusi desa digital masih belum merata. Visualisasi ini juga berfungsi sebagai dasar untuk validasi hasil model klasifikasi yang akan dikembangkan pada tahapan berikutnya.

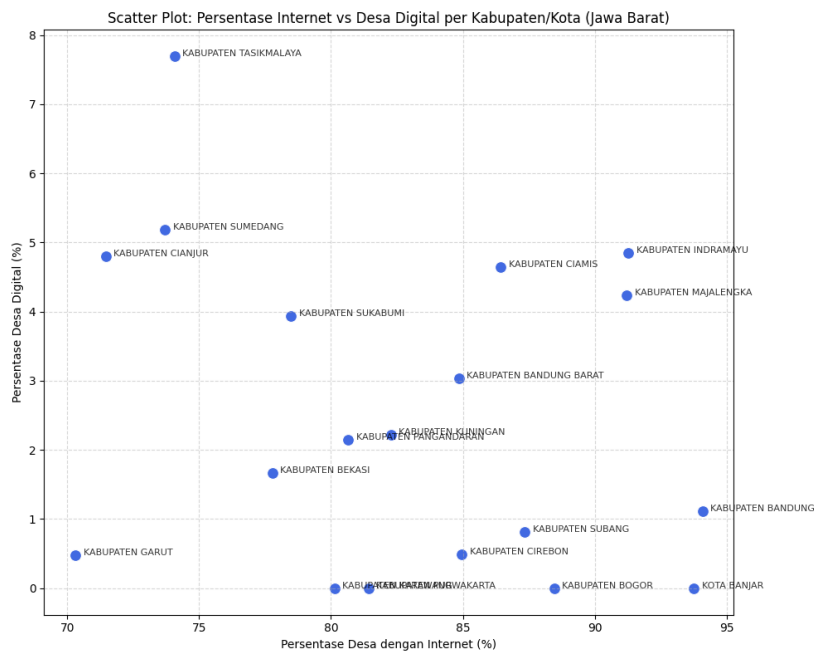


Gambar 3. Distribusi persentase desa digital per kabupaten/kota



Untuk mengevaluasi hubungan antara tingkat ketersediaan akses internet dengan persentase desa digital, digunakan visualisasi dalam bentuk *scatter plot* yang ditampilkan pada Gambar 4. Berdasarkan Gambar 4, tidak terdapat hubungan linier yang kuat antara dua variabel tersebut, misalnya Kabupaten Tasikmalaya memiliki persentase desa digital tertinggi (7.69%) namun justru memiliki akses internet yang relatif cukup rendah (74.07%) jika dibandingkan dengan kabupaten/kota lainnya. Hal ini menunjukkan bahwa status desa digital tidak selalu mencerminkan ketersediaan infrastruktur internet. Sebaliknya, wilayah dengan akses internet tinggi seperti Kabupaten Bandung, Kota Banjar, dan Kabupaten Bogor memiliki akses internet di atas 90% namun menunjukkan persentase desa digital yang rendah. Pola persebaran titik pada grafik juga tampak menyebar tanpa kecenderungan linier yang kuat, menunjukkan hubungan non-deterministik antara variabel. Dalam konteks kebijakan, visualisasi ini dapat digunakan sebagai alat untuk mengidentifikasi outlier atau daerah dengan performa tidak seimbang antara kesiapan infrastruktur dan capaian program digitalisasi desa.

Kabupaten Indramayu merupakan contoh yang relatif seimbang, dengan persentase internet tinggi (~92%) dan persentase desa digital yang juga tinggi (~4,8%). Ini menunjukkan bahwa ada wilayah yang mulai menunjukkan konsistensi antara kesiapan infrastruktur dan status digitalisasi formal. *Scatter plot* ini memperkuat argumen bahwa klasifikasi kesiapan digital desa tidak bisa semata-mata didasarkan pada keberadaan akses internet. Diperlukan pendekatan yang lebih menyeluruh dengan mempertimbangkan faktor administratif, sosial, dan kesiapan kelembagaan. Visualisasi ini juga mengindikasikan pentingnya penguatan integrasi antara pembangunan infrastruktur dan kebijakan digitalisasi formal.



Gambar 4. Scatter plot hubungan antara akses internet dengan persentase desa digital

3.3 Evaluasi Model

Sebelum dilakukan pemodelan klasifikasi, ditetapkan terlebih dahulu kategorisasi tingkat kesiapan digital kabupaten/kota berdasarkan nilai persentase desa digital. Penentuan ambang batas dilakukan dengan pendekatan kuartil atas (persentil ke-90) yaitu persentase desa digital > 5% untuk mengidentifikasi wilayah dengan dua kelas kategorisasi seperti yang dapat dilihat pada Tabel 3. Pemilihan persentil ke-90 sebagai ambang batas klasifikasi telah digunakan dalam berbagai penelitian terdahulu, khususnya untuk mengidentifikasi entitas dengan performa tertinggi dalam distribusi data. Sebagai contoh, Bornmann *et al.* (2013) [32] memanfaatkan threshold kuartil ke-90 untuk mengelompokkan publikasi ilmiah ke dalam kategori top 10% berdasarkan jumlah sitasi, sebagai indikator kinerja akademik yang unggul. Demikian pula, Teixeira-Pinto dan Normand (2008) [33] menerapkan pendekatan serupa dalam konteks evaluasi kinerja rumah sakit, di mana unit-unit layanan kesehatan dengan skor di atas persentil ke-90 diklasifikasikan sebagai berkinerja tinggi. Penggunaan ambang batas ini dinilai efektif dalam membedakan kelompok entitas dengan karakteristik istimewa di antara populasi yang diamati.

Tabel 3. Distribusi antar kelas

Nomor	Kelas	Jumlah Kabupaten/Kota
1	Tinggi	2
2	Rendah	17

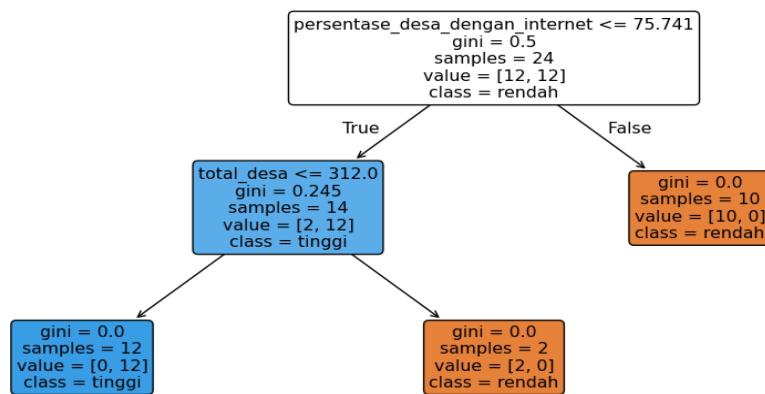


Alasan dipilihnya pendekatan kuartil ke-90 adalah untuk memastikan bahwa kategori ‘tinggi’ benar-benar mencerminkan outlier positif, atau wilayah yang telah sangat siap secara digital. Pemilihan ambang batas ini juga konsisten dengan prinsip konservatif dalam klasifikasi, yaitu lebih baik mengelompokkan wilayah yang tidak terlalu siap ke dalam kategori ‘rendah’. Dengan metode ini, terdapat dua kabupaten teridentifikasi sebagai *readiness* tinggi yaitu Kabupaten Tasikmalaya dan Kabupaten Sumedang.

Pemodelan dilakukan menggunakan algoritma *Decision Tree Classifier* dari pustaka scikit-learn. Parameter yang digunakan untuk konstruksi model yaitu: 1) `max_depth = 2`, menggambarkan batas maksimal kedalaman pohon dibatasi hingga 2 tingkat. Pembatasan ini dilakukan untuk menjaga agar model tetap sederhana dan mudah diinterpretasikan, dan; 2) `class_weight = ‘balanced’`, yang digunakan karena distribusi kelas pada target variabel tidak seimbang. Parameter ini akan otomatis menyesuaikan bobot setiap kelas secara proporsional terhadap kebalikannya dalam data latih.

Hasil klasifikasi berupa pohon keputusan dapat dilihat pada Gambar 5. Node akar pada pohon keputusan merupakan variabel `persentase_desa_dengan_internet`. Jika suatu wilayah memiliki akses internet di atas 75%, maka akan dikategorikan *readiness*-nya rendah. Hal ini menunjukkan infrastruktur saja tidak cukup untuk menunjukkan kesiapan digital yang tinggi. Pada cabang sebelah kiri, terdapat pemisahan berdasarkan variabel `total_desa`, dengan kabupaten/kota yang memiliki `total_desa` kurang dari 312 dikategorikan sebagai kabupaten/kota dengan *readiness* tinggi.

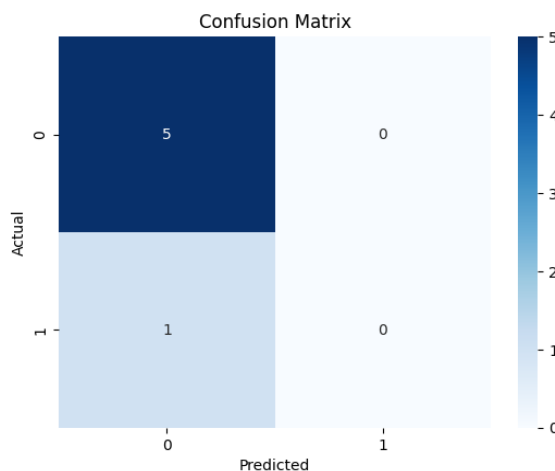
Pohon Keputusan: Klasifikasi Digital Readiness Wilayah



Gambar 5. Pohon keputusan klasifikasi *digital readiness*

Pohon keputusan ini memberikan *insight* menarik: bahwa dalam konteks Jawa Barat, akses internet yang tinggi justru tidak menjamin *readiness* tinggi. Interpretasi ini sejalan dengan temuan sebelumnya bahwa ada kabupaten dengan konektivitas tinggi tetapi adopsi digital rendah. Sementara itu, jumlah desa menjadi faktor penentu kedua, yang mungkin menggambarkan bahwa kabupaten kecil lebih adaptif dan terorganisir dalam mengimplementasikan desa digital.

Hasil *confusion matrix* model klasifikasi dapat dilihat pada Gambar 6. Meskipun akurasi keseluruhan model mencapai 83%, performa pada kelas ‘tinggi’ (1) masih sangat rendah. Untuk kelas ‘rendah’ memiliki nilai *precision* 0.83, *recall* 1.00 dan *F1-score* 0.91. Sementara kelas ‘tinggi’ memiliki nilai *precision*, *recall*, dan *F1-score* sebesar 0.00. Hal ini dikarenakan model gagal mengenali satu-satunya observasi kelas tinggi dalam data uji, walaupun sudah diberi satu observasi pada data latih. Ketidakseimbangan kelas pada data ini sangat signifikan, hanya terdapat 1 observasi kelas tinggi pada data latih dan data uji, sehingga menyebabkan model sulit belajar pola representatif dari kelas minoritas meskipun telah dicoba penanganan dengan teknik seperti SMOTE. Selain itu, ukuran dataset yang kecil juga membatasi kemampuan model untuk melakukan generalisasi.

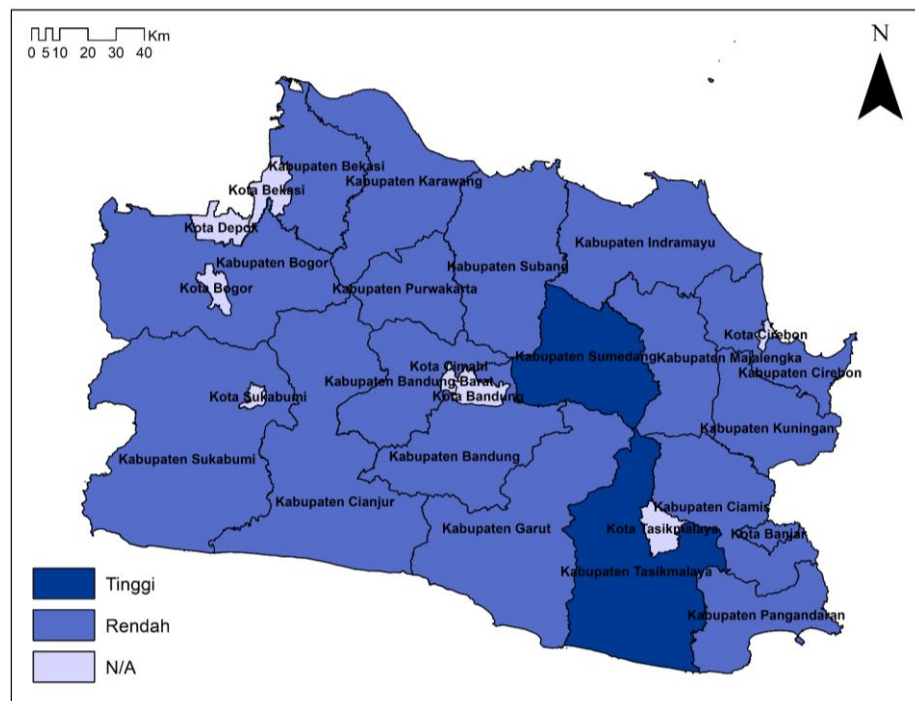


Gambar 6. *Confusion matrix* model klasifikasi

Kinerja yang rendah untuk kelas ‘tinggi’ tidak dapat langsung ditafsirkan sebagai kelemahan model, tetapi sebagai cerminan dari tantangan klasifikasi pada data tidak seimbang. Hal ini juga memperkuat argumen bahwa pengumpulan data tambahan sangat diperlukan untuk meningkatkan ketahanan dan validitas model.

3.4 Pembahasan

Visualisasi spasial hasil klasifikasi tingkat kesiapan digital (*digital readiness*) kabupaten/kota di Provinsi Jawa Barat disajikan pada Gambar 7. Berdasarkan hasil klasifikasi yang dibangun menggunakan algoritma *Decision Tree*, wilayah dikategorikan ke dalam dua kelas utama: ‘tinggi’ dan ‘rendah’. Kategori ini didasarkan pada ambang batas kuartil ke-90 dari nilai persentase desa digital, yaitu sebesar 5%. Dengan ambang ini, hanya dua kabupaten yang termasuk dalam kategori *readiness* ‘tinggi’, yakni Kabupaten Tasikmalaya dan Kabupaten Sumedang, yang ditunjukkan dengan warna biru tua pada peta. Sebaliknya, sebagian besar kabupaten lainnya diklasifikasikan sebagai *readiness* ‘rendah’ (warna biru muda), sementara beberapa wilayah kota ditandai sebagai tidak tersedia data (N/A) karena keterbatasan informasi yang dapat diperoleh selama tahap akuisisi dan pra-pemrosesan data.



Gambar 7. Visualisasi Hasil Klasifikasi

Hasil visualisasi ini konsisten dengan analisis eksploratif sebelumnya, di mana Kabupaten Tasikmalaya dan Kabupaten Sumedang menempati urutan teratas dalam hal persentase desa digital. Sementara itu, wilayah dengan ketersediaan internet tinggi namun rendah dalam adopsi desa digital tetap tergolong *readiness* rendah, yang secara empirik mengindikasikan bahwa konektivitas internet belum cukup untuk mendorong transformasi digital secara menyeluruh. Distribusi spasial ini menggarisbawahi ketimpangan kesiapan digital antar wilayah di Jawa Barat. Wilayah dengan jumlah desa yang relatif sedikit namun proporsi desa digital tinggi cenderung diklasifikasikan sebagai siap secara digital, sebagaimana tercermin dari pohon keputusan hasil pemodelan. Temuan ini penting dalam konteks perumusan kebijakan pengembangan desa digital, karena dapat digunakan untuk memprioritaskan wilayah yang memerlukan intervensi lebih lanjut, baik dalam bentuk infrastruktur, literasi digital, maupun pendampingan kelembagaan.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi tingkat kesiapan digital desa di Provinsi Jawa Barat dengan menggunakan algoritma *Decision Tree*. Dengan memanfaatkan data sekunder yang bersumber dari portal Open Data Jabar, hasil analisis model menunjukkan bahwa persentase desa dengan internet dan jumlah desa digital merupakan dua indikator utama yang mempengaruhi klasifikasi *digital readiness* suatu wilayah. Model menghasilkan akurasi klasifikasi sebesar 83%, meskipun performa pada kelas *readiness* ‘tinggi’ masih rendah akibat ketidakseimbangan jumlah data antar kelas. Temuan penting dari penelitian ini adalah bahwa ketersediaan infrastruktur internet yang tinggi tidak serta merta berkorelasi dengan tingginya status digitalisasi desa. Wilayah seperti Kabupaten Bandung dan Kota Banjar menunjukkan tingkat konektivitas internet yang tinggi, tetapi tidak disertai dengan proporsi desa digital yang memadai. Sebaliknya, daerah seperti Kabupaten Sumedang dan Tasikmalaya, meskipun infrastrukturnya tidak paling unggul, berhasil menunjukkan kesiapan digital tinggi karena proporsi desa mandiri yang besar. Selain itu, analisis visualisasi spasial memperkuat kesimpulan bahwa ketimpangan *digital readiness* antarwilayah cukup signifikan. Beberapa wilayah



menunjukkan tingginya akses internet namun belum mampu mengonversi potensi tersebut menjadi status digital yang formal. Hal ini menunjukkan perlunya pemahaman yang lebih komprehensif terhadap dinamika pembangunan digital desa, termasuk interaksi antara aspek infrastruktur, sosial, dan kelembagaan yang saling terkait dan tidak dapat dipisahkan dalam menciptakan ekosistem digital yang berkelanjutan di tingkat lokal.

REFERENCES

- [1] Kementerian Komunikasi dan Digital, “Menghubungkan Pelosok Indonesia Menuju Masa Depan.” Accessed: Aug. 08, 2025. [Online]. Available: <https://www.komdigi.go.id/berita/pemerintahan-digital/detail/menghubungkan-pelosok-indonesia-menuju-masa-depan>
- [2] Good Stats, “Survei APJII: 82,6% Penduduk Daerah Tertinggal Sudah Memiliki Akses Internet.” [Online]. Available: <https://goodstats.id/article/survei-apjii-82-6-penduduk-daerah-tertinggal-sudah-memiliki-akses-terhadap-internet-0Ic8n>
- [3] Kominfo, *Status Literasi Digital di Indonesia 2021*. 2022. [Online]. Available: https://cdn1.katadata.co.id/media/microsites/litdik/Status_Literasi_Digital_diIndonesia%20_2021_190122.pdf
- [4] R. Chinoracky, N. Stalmasekova, R. Madlenak, and L. Madlenakova, “Are Nations Ready for Digital Transformation? A Macroeconomic Perspective Through the Lens of Education Quality,” *Economies*, vol. 13, no. 6, p. 152, May 2025, doi: 10.3390/economies13060152.
- [5] X. Duanmu, J. Yu, X. Yuan, and X. Zhang, “How Does Digital Infrastructure Mitigate Urban–Rural Disparities?,” *Sustainability*, vol. 17, no. 4, p. 1561, Feb. 2025, doi: 10.3390/su17041561.
- [6] F. F. Maimuna, N. A. F. Roroa, M. Misrah, O. Oktavianty, and A. Agit, “Transformasi digital dalam kewirausahaan: Analisis faktor penghambat dan pendorong perkembangan ekonomi digital,” in *Prosiding Seminar Nasional Pembangunan Ekonomi Berkelanjutan Dan Riset Ilmu Sosial*, 2024, pp. 187–198.
- [7] R. Jayanthi and A. Dinaseviani, “Kesenjangan Digital dan Solusi yang Diterapkan di Indonesia Selama Pandemi COVID-19,” *J. IPTEKKOM J. Ilmu Pengetah. Teknol. Inf.*, vol. 24, no. 2, pp. 187–200, Dec. 2022, doi: 10.17933/iptekkom.24.2.2022.187-200.
- [8] Y. D. Hadiyat and others, “Kesenjangan Digital di Indonesia (Studi Kasus di Kabupaten Wakatobi),” *J. Pekommas*, vol. 17, no. 2, pp. 81–90, 2014, doi: 10.30818/jpkm.2014.1170203.
- [9] E. Firmansyah and M. A. Helmiawan, “Pengukuran Kesiapan Transformasi Digital Desa Kaduwulung Menuju Desa Cerdas Berbasis SNI ISO 37122: 2019 Melalui Pemetaan Data Desa,” *Infomans J. Ilmu-Ilmu Inform. Dan Manaj.*, vol. 19, no. 1, 2025.
- [10] A. Koswara, “Digitalisasi Ekonomi di Pedesaan: Mengkaji Kesenjangan Infrastruktur Digital di Indonesia,” *J. Al Azhar Indones. Seri Ilmu Sos.*, vol. 5, no. 3, p. 180, Nov. 2024, doi: 10.36722/jaiss.v5i3.3407.
- [11] Badan Pusat Statistik, “Jumlah Desa/Kelurahan Menurut Provinsi, 2024.” Feb. 26, 2025. Accessed: Aug. 08, 2025. [Online]. Available: <https://www.bps.go.id/id/statistics-table/3/YkVWWFUYnTJTVTloVGpCeFdFVTNaMk5wUzFaUFFUMDkjMw==/jumlah-desa-sup-1--sup--kelurahan-menurut-provinsi--2024.html?year=2024>
- [12] U. Madjid, M. W. Kawuryan, A. Averus, and T. Triyanto, “Communication in Digital-Based Public Services in Regional Government of West Java Province,” *Transform. J. Manaj. Pemerintah.*, pp. 48–61, July 2024, doi: 10.33701/jtp.v16i1.3778.
- [13] M. A. A. Hilmi, M. Mustamiin, A. Nagi, A. Suheryadi, and F. P. B. Muhammad, “Uji Performa dan Website Responsiveness Institusi dan Smart City se-Jawa Barat,” *ArXiv Prepr. ArXiv191213346*, 2019, doi: 10.48550/arXiv.1912.13346.
- [14] F. Pirola, C. Cimini, and R. Pinto, “Digital readiness assessment of Italian SMEs: a case-study research,” *J. Manuf. Technol. Manag.*, vol. 31, no. 5, pp. 1045–1083, Nov. 2019, doi: 10.1108/JMTM-09-2018-0305.
- [15] M. V. Tsurkan, S. M. Mironova, and N. V. Pilipchuk, “Digital transformation of project implementation monitoring in the regional public management,” in *Proceedings of the 1st International Scientific Conference “Modern Management Trends and the Digital Economy: from Regional Development to Global Economic Growth” (MTDE 2019)*, Yekaterinburg, Russia: Atlantis Press, 2019. doi: 10.2991/mtde-19.2019.15.
- [16] R. Musfikar, U. Rizqina, and Y. Yusran, “Analisis Kesiapan Desa Menuju Smart Village Pada Kecamatan Indrapuri Aceh Besar,” *J. Infomedia*, vol. 7, no. 2, p. 86, Dec. 2022, doi: 10.30811/jim.v7i2.3025.
- [17] Aditianto, Murhadi, and H. M. Jumasa, “Analisis Tingkat Kesiapan Pemanfaatan Sistem Informasi Desa Di Desa Kedungpomahan Wetan Purworejo,” *INTEK J. Inform. Dan Teknol. Inf.*, vol. 7, no. 2, pp. 18–25, Nov. 2024, doi: 10.37729/intek.v7i2.5321.
- [18] R. Mayasari, N. Heryana, and A. Hananto, “Village Government Readiness Toward the Adoption of E-Government,” *Buana Inf. Technol. Comput. Sci. BIT CS*, vol. 4, no. 1, pp. 6–10, Jan. 2023, doi: 10.36805/bit-cs.v4i1.3241.
- [19] S. Subhan and A. B. Pratama, “Kesiapan Desa dalam Penerapan E-Government (Studi Perbandingan di Desa Semaya dan Desa Pegiringan Kabupaten Pematang),” *Sawala J. Adm. Negara*, vol. 7, no. 2, pp. 161–175, Dec. 2019, doi: 10.30656/sawala.v7i2.1542.
- [20] Z. Sitorus, E. Hariyanto, and F. Kurniawan, “IMPLEMENTASI MACHINE LEARNING SEBAGAI SISTEM PEMETAAN DAERAH RAWAN BANJIR DESA PAHLAWAN,” *Penerbit Tahta Media*, 2023.
- [21] R. Regen and H. Setiawan, “Advancing Cardiovascular Risk Prediction: A Review of Machine Learning Models and Their Clinical Potential,” *J. Electr. Technol. UMY*, vol. 8, no. 2, pp. 51–59, 2024, doi: 10.18196/jet.v8i2.25208.
- [22] A. Priyam, G. R. Abhijeeta, A. Rathee, and S. Srivastava, “Comparative analysis of decision tree classification algorithms,” *Int. J. Curr. Eng. Technol.*, vol. 3, no. 2, pp. 334–337, 2013.
- [23] S. R. Pingali, S. Singha, S. Arunachalam, and K. Pedada, “Digital readiness of small and medium enterprises in emerging markets: The construct, propositions, measurement, and implications,” *J. Bus. Res.*, vol. 164, p. 113973, 2023, doi: <https://doi.org/10.1016/j.jbusres.2023.113973>.
- [24] F. Michelotto and L. A. Joia, “Organizational Digital Transformation Readiness: An Exploratory Investigation,” *J. Theor. Appl. Electron. Commer. Res.*, vol. 19, no. 4, pp. 3283–3304, Nov. 2024, doi: 10.3390/jtaer19040159.
- [25] M. T. Islam, A. A.-T. Hasan, and A. Faysal, “Factors influencing the readiness of rural people to use Union Digital Center services in southern Bangladesh: a behavioral intention approach,” *J. Electron. Bus. Digit. Econ.*, 2025, doi: 10.1108/JEBDE-09-2024-0036.



- [26] F. Sulistyowati, H. S. Tyas, M. C. R. Dibyorini, C. Puspitosari, and others, "Pemanfaatan Sistem Informasi Desa (SID) untuk Mewujudkan Smart Village di Kalurahan Panggungharjo DIY (Utilization of Sistem Informasi Desa (SID) to Realize Smart Village in Kalurahan Panggungharjo, Sewon, Bantul, DI Yogyakarta)," *J. IPTEKKOM J. Ilmu Pengetah. Teknol. Inf.*, vol. 23, no. 2, pp. 213–226, 2021, doi: 10.17933/iptekkom.23.2.2021.213-226.
- [27] A. Premana, H. Sucipto, and A. Widiatoro, "Pengembangan desa berbasis smart village (Studi smart governance pada pelayanan prima Desa Tegalreja)," *J. Ilm. Pengabd. Dan Inov.*, vol. 1, no. 1, pp. 43–54, 2022.
- [28] A. Zheng and A. Casari, *Feature Engineering for Machine Learning*. O'Reilly, 2018. [Online]. Available: <https://www.oreilly.com/library/view/feature-engineering-for/9781491953235/>
- [29] I. Setiawan, R. Fina Antika Cahyani, and I. Sadida, "EXPLORING COMPLEX DECISION TREES: UNVEILING DATA PATTERNS AND OPTIMAL PREDICTIVE POWER," *J. Innov. Future Technol. IFTECH*, vol. 5, no. 2, pp. 112–123, Sept. 2023, doi: 10.47080/iftech.v5i2.2829.
- [30] W. Musu, A. Ibrahim, and H. Heriadi, "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4. 5," in *SISITI: Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, 2021, pp. 186–195. doi: 10.36774/sisiti.v10i1.802.
- [31] M. FADLI and R. A. Saputra, "Klasifikasi dan evaluasi performa model Random Forest untuk prediksi stroke," *J. Tek.*, vol. 12, no. 2, 2023, doi: 10.31000/jt.v12i2.9099.
- [32] L. Bornmann, L. Leydesdorff, and R. Mutz, "The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits," *J. Informetr.*, vol. 7, no. 1, pp. 158–165, Jan. 2013, doi: 10.1016/j.joi.2012.10.001.
- [33] A. Teixeira-Pinto and S. T. Normand, "Statistical methodology for classifying units on the basis of multiple-related measures," *Stat. Med.*, vol. 27, no. 9, pp. 1329–1350, Apr. 2008, doi: 10.1002/sim.3187.