SKRIPSI Kamia Pertiwi Sistem Informasi.pdf

WORD COUNT 13895
TIME SUBMITTED 01-SEP-2026 06:09PM
PAPER ID 124054764

PREDIKSI RISIKO KETERGANTUNGAN SMARTPHONE PADA ANAK USIA DINI MENGGUNAKAN ALGORITMA RANDOM FOREST

SKRIPSI

DISUSUN OLEH :

KAMIA PERTIWI

403221010005

PRODI SISTEM INFORMASI

FAKULTAS TEKNIK DAN ILMU KOMPUTER

UNIVERSITAS ISLAM INDRAGIRI

2026

PREDIKSI RISIKO KETERGANTUNGAN SMARTPHONE PADA ANAK USIA DINI MENGGUNAKAN ALGORITMA RANDOM FOREST

PREDICTING THE RISK OF SMARTPHONE ADDICTION IN EARLY CHILDHOOD USING THE RANDOM FOREST ALGORITHM

Diajukan untuk memenuhi salah satu syarat

Memperoleh derajat Serjana S1

DI SUSUN OLEH :

KAMIA PERTIWI

403221010005

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK DAN ILMU KOMPUTER

UNIVERSITAS ISLAM INDRAGIRI

2026

PERSETUJUAN

PREDIKSI RISIKO KETERGANTUNGAN SMARTPHONE PADA ANAK USIA DINI MENGGUNAKAN ALGORITMA RANDOM FOREST

yang dipersiapkan dan disusun oleh

KAMIA PERTIWI
403221010005

telah disetujui oleh Dosen Pembimbing Tugas Akhir

pada tanggal 18 Desember 2025

PENGESAHAN

PREDIKSI RISIKO KETERGANTUNGAN SMARTPHONE PADA ANAK USIA DINI MENGGUNAKAN ALGORITMA RANDOM FOREST

yang dipersiapkan dan disusun oleh

KAMIA PERTIWI
403221010005

telah disetujui oleh Dosen Pembimbing Tugas Akhir

pada tanggal **18 Desember 2025**

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggung jawab saya pribadi.

Tembilahan, 2026

KAMIA PERTIWI

KITA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT, Tuhan yang Maha Membuka segala pintu kemudahan dan Maha Melimpahkan cahaya ilmu kepada setiap hamba yang berusaha. Dengan rahmat, kasih, dan ketetapan-Nya, proposal penelitian berjudul "Prediksi Risiko Ketergantungan Smartphone pada Anak Usia Dini Menggunakan Algoritma *Random Forest*" ini akhirnya dapat tersusun dengan baik. Proses yang ditempuh mulai dari menelusuri literatur, merumuskan gagasan, mengolah konsep, hingga menyusunnya menjadi sebuah karya ilmiah merupakan perjalanan yang dipenuhi pelajaran, hambatan, renungan, serta ketekunan yang membentuk pemahaman baru bagi penulis.

Dengan penuh kesadaran, penulis memahami bahwa proposal ini tidak mungkin hadir tanpa cahaya, arahan, dan dorongan yang diberikan oleh berbagai pihak yang telah menjadi bagian penting dalam proses penyusunannya. Oleh karena itu, dengan segala kerendahan hati dan rasa hormat, penulis menyampaikan terima kasih yang sedalam-dalamnya kepada:

1. Bapak Dr. H. Najamuddin, Lc., M.A, selaku Rektor Universitas Islam Indragiri, yang melalui kebijakan, visi, dan komitmennya terhadap dunia pendidikan telah membuka jalan bagi mahasiswa untuk berkembang, berkarya, serta berkontribusi melalui penelitian. Dukungan beliau merupakan landasan kuat yang memungkinkan penelitian ini dapat terlaksana.
2. Bapak Dr. Bayu Rianto, S.SI., M.Kom, selaku Dekan Fakultas Teknik dan Ilmu Komputer, yang senantiasa memberikan dukungan akademik dan lingkungan belajar yang kondusif, sehingga proses penelitian dapat berjalan secara terarah, terstruktur, dan penuh motivasi.

3. Ibu Fitri Yunita, S.SI., M.Kom, selaku Ketua Program Studi Sistem Informasi, yang tidak hanya memberikan arahan akademik, tetapi juga menghadirkan perhatian dan semangat yang menguatkan mahasiswa dalam meniti perjalanan ilmiah.
4. Bapak Usman, ST., M.Kom, selaku pembimbing utama sekaligus pembimbing akademik, yang dengan ketelatenan dan kesabarannya memberikan bimbingan yang sangat berarti. Setiap saran, koreksi, dan evaluasi beliau menjadi pijakan penting yang memperkaya kualitas penelitian ini dari segi konsep, teknis, hingga akurasi prediksi. Penulis mengakui bahwa banyak pembelajaran berharga diperoleh dari proses diskusi dan bimbingan beliau.
5. Bapak Ilyas, S.Kom., M.Kom, selaku pembimbing pendamping, atas segala petunjuk dan pandangan kritis yang diberikan. Beliau membantu penulis melihat permasalahan penelitian dari perspektif yang lebih jernih, analitis, dan objektif. Setiap masukan beliau berkontribusi besar terhadap penyempurnaan kerangka penelitian ini.
6. Seluruh dosen, staf, dan civitas akademika Program Studi Sistem Informasi Universitas Islam Indragiri, yang telah menanamkan dasar-dasar keilmuan, membimbing dengan penuh dedikasi, serta menyediakan lingkungan akademik yang mendorong perkembangan pemikiran dan keterampilan ilmiah penulis.

   Ucapan terima kasih yang tak terhingga juga penulis sampaikan kepada
7. Kedua orang tua tercinta, Bapak Darmawi dan Ibu Vivi Susanti, serta adik tersayang, Ridho Darfiansyah, atas segala doa, kasih sayang, dukungan, dan

pengorbanan yang selalu menjadi sumber kekuatan penulis dalam menyelesaikan pendidikan ini.

8. Teman-teman seperjuangan, khususnya Ainur Rahmadina, Khusnul Khotimah, dan Marliyah, yang selalu hadir sebagai tempat bertanya, berbagi lelah, bertukar ide, hingga menjadi *support system* yang mengisi perjalanan akademik ini dengan warna-warni kebersamaan. Terima kasih atas segala dukungan, semangat, dan kebersamaan yang telah diberikan. Kalian telah menjadi pengingat bahwa perjalanan panjang akan terasa lebih ringan ketika ditempuh bersama.

Penulis berharap proposal penelitian ini tidak hanya menjadi bagian dari pemenuhan tugas akademik, tetapi juga dapat memberi kontribusi bagi pengembangan ilmu pengetahuan, khususnya dalam memahami perilaku digital anak usia dini melalui pendekatan teknologi modern. Penulis menyadari sepenuhnya bahwa karya ini masih jauh dari sempurna. Oleh sebab itu, kritik dan saran yang membangun sangat diharapkan untuk meningkatkan kualitas penelitian di tahap selanjutnya.

Akhir kata, semoga setiap ikhtiar dalam penyusunan penelitian ini menjadi amal ilmu yang bermanfaat, membawa keberkahan, serta menjadi langkah kecil menuju perubahan yang lebih baik di dunia akademik maupun masyarakat.

Tembilahan 06 Desember 2025
Penulis

Kamia Pertiwi
403221010005

DAFTAR ISI

HALAMAN JUDUL

PERSETUJUAN

PENGESAHAN

PERNYATAAN

KITA PENGANTAR vi

DAFTAR ISI ix

DAFTAR TABEL xi

DAFTAR GAMBAR xi

INTISARI xii

ABSTRACT xiii

BAB I PENDAHULUAN 1

1.1 Latar Belakang 1

1.2 Rumusan Masalah 3

1.3 Tujuan Penelitian 3

1.4 Batasan Penelitian 4

1.5 Manfaat Penelitian 5

1.6 Sistematika Penulisan 5

BAB II TINJAUAN PUSTAKA 7

2.1 Penelitian Terdahulu 7

2.2 Ketergantungan Smartphone 19

2.3 Anak Usia Dini 21

2.4 Machine Learning 22

2.5 Algoritma Random Forest 25
2.6 Variabel Dependen 27
2.7 Variabel Independen 28
2.8 Rangkuman 30
BAB III METODE PENELITIAN 32
3.1 Kerangka Penelitian 32
3.1.1 Penjelasan Alur Penelitian 34
3.2 Model Random Forest 38
BAB IV HASIL DAN PEMBAHASAN 40
4.1 Data Penelitian 40
4.2 Preprocessing Data 42
4.3 Pembagian Data Training dan Data Testing 44
4.4 Pelatihan Model 45
4.5 Perbandingan Performa Model 57
4.6 Uji ANOVA 58
4.7 Analisis Feature Importance 60
BAB V PENUTUP 60
5.1 Kesimpulan 60
5.2 Saran 61
DAFTAR PUSTAKA

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu 8

Tabel 2. 1 Penelitian Terdahulu Lanjutan 9

Tabel 2. 1 Penelitian Terdahulu Lanjutan 10

Tabel 2. 1 Penelitian Terdahulu Lanjutan 11

Tabel 2. 1 Penelitian Terdahulu Lanjutan 12

Tabel 2. 1 Penelitian Terdahulu Lanjutan 13

Tabel 2. 1 Penelitian Terdahulu Lanjutan 14

Tabel 2. 1 Penelitian Terdahulu Lanjutan 15

Tabel 2. 1 Penelitian Terdahulu Lanjutan 16

Tabel 2. 1 Penelitian Terdahulu Lanjutan 17

Tabel 2. 1 Confusion Matrix 8

Tabel 4. 1 Distribusi Kelas Keseluruhan 44

Tabel 4. 2 Pembagian Data Training dan Data Testing 44

Tabel 4.3 Distribusi Data Training..... 44

Tabel 4. 4 Distribusi Data Testing 45

Tabel 4.5 Parameter 47

Tabel 4.6 Contoh Pohon Keputusan 47

Tabel 4.7 Contoh Prediksi Pohon Keputusan 48

Tabel 4.8 Jumlah Prediksi..... 48

Tabel 4. 9 Classification Report Random Forest 49

Tabel 4. 10 Nilai TP, FP dan FN..... 50

Tabel 4. 11 Classification Report Model SVM..... 55

Tabel 4. 12 Classification Report Model KNN 57

Tabel 4. 13 Perbandingan Performa Model 58

Tabel 4. 14 Hasil Uji ANOVA Perbandingan Model 59

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian 33

Gambar 3.2 Random Forest 38

Gambar 4.1 Data Asli 43

Gambar 4.2 Data Asli 43

Gambar 4.3 Data Asli 43

Gambar 4.4 Data Asli 43

Gambar 4.6 Distribusi Data Training dan Data Testing 45

Gambar 4.7 Distribusi Kelas Sebelum dan Setelah Penggabungan 46

Gambar 4.8 Confusion Matrix Random Forest 49

Gambar 4.9 Confusion Matrix SVM 54

Gambar 4.10 Confusion Matrix KNN 56

Gambar 4.11 Boxplot Hasil Uji ANOVA 59

Gambar 4.12 Rata-Rata Akurasi Ketiga Model 60

Gambar 4.13 Diagram Feature Importance 60

INTISARI

Penggunaan smartphone pada anak usia dini yang semakin meningkat berpotensi menimbulkan risiko ketergantungan sehingga diperlukan metode untuk memprediksi tingkat risikonya. Penelitian ini bertujuan untuk memprediksi risiko ketergantungan smartphone pada anak usia dini menggunakan algoritma Random Forest. Penelitian menggunakan metode kuantitatif dengan data primer yang diperoleh melalui kuesioner kepada orang tua yang memiliki anak usia 1–6 tahun. Data yang digunakan berjumlah 347 responden dengan enam variabel independen, yaitu umur, durasi penggunaan smartphone, tujuan penggunaan smartphone, pola pengawasan orang tua, frekuensi penggunaan smartphone, dan usia pertama kali mengenal smartphone. Tahapan penelitian meliputi preprocessing data, pembagian data, pembangunan model Random Forest, serta perbandingan dengan algoritma Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN). Evaluasi model dilakukan menggunakan accuracy, precision, recall, F1-score, dan confusion matrix. Hasil penelitian menunjukkan bahwa Random Forest memberikan performa terbaik dengan accuracy 87%, precision 61%, recall 55%, dan F1-score 56%, lebih tinggi dibandingkan SVM dan KNN. Selain itu, pola pengawasan orang tua merupakan variabel yang paling berpengaruh terhadap risiko ketergantungan smartphone pada anak usia dini.

Kata kunci: Anak Usia Dini, Ketergantungan Smartphone, Machine Learning, Prediksi, Random Forest.

ABSTRACT

The increasing use of smartphones among early childhood children may lead to addiction, making an accurate prediction model necessary. This study aims to predict the risk of smartphone addiction in early childhood using the Random Forest algorithm. A quantitative method was employed using primary data collected through questionnaires from parents of children aged 1–6 years. The dataset consisted of 347 respondents with six independent variables: age, smartphone usage duration, smartphone usage purpose, parental supervision, smartphone usage frequency, and age at first smartphone exposure. The research process included data preprocessing, data splitting, Random Forest model development, and comparison with Support Vector Machine (SVM) and K-Nearest Neighbor (KNN). Model performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrix. The results showed that Random Forest achieved the best performance with an accuracy of 87%, precision of 61%, recall of 55%, and F1-score of 56%, outperforming SVM and KNN. Feature importance analysis also indicated that parental supervision was the most influential factor in predicting smartphone addiction risk among early childhood children.

Keywords: Early Childhood, Machine Learning, Prediction, Random Forest, Smartphone Addicti

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital saat ini membawa dampak besar terhadap kehidupan manusia, termasuk dalam aspek komunikasi, hiburan, maupun pendidikan. Salah satu bentuk teknologi yang paling banyak digunakan adalah smartphone. Tidak hanya orang dewasa, anak usia dini pun kini sudah terbiasa menggunakan smartphone, baik untuk menonton video,bermain permainan interaktif, maupun mengakses aplikasi pembelajaran. Fenomena ini menunjukkan adanya perubahan pola aktivitas anak di era digital.[1]

Meskipun penggunaan smartphone dapat memberikan manfaat tertentu, terutama dalam hal kemudahan akses informasi dan media pembelajaran, penggunaan yang berlebihan pada anak usia dini berpotensi menimbulkan masalah serius. Anak dapat mengalami gejala kecanduan atau ketergantungan, seperti sulit lepas dari smartphone, mudah marah ketika dibatasi, hingga berkurangnya minat pada aktivitas sosial secara langsung. Hal ini berpotensi mengganggu perkembangan sosial, emosional, dan kognitif anak. [2]

Berdasarkan data BPS (2023), jumlah pengguna gadget untuk anak usia dini di Indonesia mencapai 33,44%, dengan rincian 25,5% pada anak berusia 0–4 tahun dan 52,76% pada anak berusia 5–6 tahun. Kondisi ini menunjukkan potensi yang cukup besar terhadap munculnya kecanduan gadget pada anak. Fenomena ini diperkuat oleh survei Komisi Perlindungan Anak Indonesia (KPAI, 2022) yang menyebutkan bahwa lebih dari 71,3% anak usia sekolah memiliki gadget dan menggunakannya dalam porsi waktu yang cukup lama setiap hari. Selain itu,

sebanyak 79% anak diperbolehkan bermain gadget di luar keperluan belajar. Situasi tersebut sering kali tidak disadari oleh orang tua, sehingga anak-anak yang menghabiskan terlalu banyak waktu di depan layar berisiko mengalami berbagai masalah, mulai dari gangguan tidur, penurunan prestasi akademik, hingga masalah sosial dan emosional.[3]

Beberapa penelitian sebelumnya telah menyoroti dampak penggunaan smartphone terhadap perilaku anak dan remaja, namun sebagian besar kajian masih bersifat deskriptif dan berfokus pada aspek pskilogis. Sementara itu, penelitian yang mengintegrasikan pendekatan machine learning dalam memprediksi risiko ketergantungan smartphone pada anak usia dini masih sangat terbatas. Celah penelitian ini penting untuk diisi, mengingat metode prediktif berbasis data dapat membantu memberikan gambaran lebih objektif terkait tingkat risiko yang dihadapi anak.

Dalam konteks ini, algoritma Random Forest dipilih sebagai metode prediksi karena memiliki keunggulan dalam menangani data yang kompleks, menghasilkan tingkat akurasi yang baik berdasarkan temuan penelitian sebelumnya, serta mampu mengidentifikasi variable-variabel yang paling berpengaruh. Dengan demikian, penerapan algoritma ini diharapkan mampu menghasilkan model prediksi yang dapat membantu orang tua, pendidik, maupun pihak terkait dalam mengantisipasi dan mengurangi risiko ketergantungan smartphone pada anak sejak dini.[4]

Penelitian ini dilakukan dengan memanfaatkan data yang diperoleh melalui kuesioner serta dataset pendukung. Data tersebut digunakan untuk

mengidentifikasi pola penggunaan smartphone pada anak usia dini dan factor-faktor yang memengaruhi tingkat ketergantungan.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Diperlukan penerapan algoritma Random Forest untuk membangun model prediksi dalam mengidentifikasi risiko ketergantungan smartphone pada anak usia dini sebagai dasar upaya pencegahan sejak dini.
2. Belum tersedia model prediksi untuk mengukur risiko ketergantungan smartphone pada anak usia dini secara objektif.
3. Belum diketahui faktor yang paling berpengaruh dalam memprediksi risiko ketergantungan smartphone pada anak usia dini

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, tujuan dari penelitian ini adalah:

1. Membangun model prediksi risiko ketergantungan smartphone pada anak usia dini menggunakan algoritma random forest sebagai algoritma utama.
2. Menguji performa model *Random Forest* dan membandingkannya dengan algoritma (SVM dan KNN) sebagai algoritma pembanding menggunakan uji ANOVA.
3. Mengetahui faktor yang paling berpengaruh dalam memprediksi risiko ketergantungan smartphone pada anak usia dini berdasarkan hasil analisis *feature importance* pada algoritma Random Forest.

1.4 Batasan Penelitian

Untuk meminimalisir perluasan dalam pembahasan penelitian, adapun batasan-batasan masalah yang diberikan sebagai berikut:

1. Data yang digunakan adalah data primer yang diperoleh dari kuesioner dengan orang tua atau wali anak usia umur 1-6 tahun sebagai responden.
2. Variabel yang digunakan antara lain tingkat risiko ketergantungan smartphone pada anak uisa dini(variabel dependen Y), umur, durasi penggunaan smarthphone perhari, frekuensi, usia awal kenal smartphone. tujuan penggunaan smartphone, pola pengawasan orang tua, (variabel independen X).
3. Penelitian ini berfokus pada penerapan algoritma *Random Forest* sebagai algoritma utama, sedangkan SVM daan KNN hanya digunakan sebagai algoritma pembanding untuk mendukung analisis performa model.
4. Uji ANOVA digunakan sebagai uji statistik pendukung untuk melihat perbedaan signifikan performa antar model.
5. Hasil penelitian berupa model prediksi risiko ketergantungan smartphone, bukan rekomendasi intervensi psikologis atau klinis.
6. Penelitian ini dibatasi pada penggunaan data sebanyak 347 responden yang diperoleh melalui penyebaran kuesioner kepada orang tua atau wali anak usia 1–6 tahun.
7. Penelitian ini dibatasi pada orang tua yang memiliki anak usia 1–6 tahun yang berdomisili di Desa Harapan Tani, Kecamatan Kempas, Kabupaten Indragiri Hilir. Pengambilan data dilakukan pada wilayah yang berhasil dijangkau oleh

peneliti sehingga cakupan responden tidak mewakili seluruh wilayah Desa Harapan Tani.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik dari sisi teoritis maupun praktis. Dari sisi teoritis, peneltian ini berkontribusi dalam memperkaya kajian akademik mengenai penerapan algoritma *machine learning* di bidang perilaku digital anak usia dini, khususnya dalam konteks prediksi risiko ketergantungan smartphone. Integrasi faktor demografis, psikologis, lingkungan keluarga, dan perilaku teknis dalam satu model prediksi juga memberikan nilai tambah bagi pengembangan penelitian serupa di masa mendatang.

Dari sisi praktis, penelitian ini diharapkan dapat menjadi acuan bagi orang tua, pendidik, dan pihak terkait dalam mengidentifikasi tingkat risiko ketergantungan smartphone pada anak usia dini secara lebih objektif dan terukur. Model prediksi yang dihasilkan diharapkan dapat membantu dalam perumusan strategi pengawasan, pembiasaan digital yang sehat, serta penyusunan kebijakan literasi digital anak yang seimbang antara aspek teknologi dan perkembangan psikososial anak di Indonesia.

1.6 Sistematika Penulisan

Agar pembahasan dalam skripsi ini lebih terstruktur dan mudah dipahami, penulisan disusun ke dalam beberapa bab dengan sistematika sebagai berikut:

BAB I PENDAHULUAN

Berisi latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan penelitian, serta sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Memuat landasan teori yang mendukung penelitian, kajian literatur terkait penggunaan smartphone pada anak usia dini, konsep ketergantungan, serta penjelasan mengenai algoritma Random Forest. Selain itu, bab ini juga mencakup penelitian-penelitian terdahulu yang relevan sebagai dasar dalam merumuskan kerangka pemikiran penelitian.

BAB III METODOLOGI PENELITIAN

Menjelaskan pendekatan penelitian yang digunakan, sumber data, teknik pengumpulan data melalui kuesioner dan dataset, variabel penelitian, tahapan pengolahan data, serta metode analisis menggunakan algoritma Random Forest.

BAB IV HASIL PENELITIAN

Menyajikan hasil pengolahan data, implementasi algoritma Random Forest, hasil evaluasi model, perbandingan dengan algoritma SVM dan KNN, hasil uji ANOVA, analisis feature importance, serta pembahasan terhadap hasil penelitian yang diperoleh.

BAB V PENUTUP

Berisi kesimpulan yang diperoleh dari hasil penelitian serta saran yang dapat dijadikan sebagai bahan pengembangan penelitian selanjutnya maupun penerapan model prediksi risiko ketergantungan smartphone pada anak usia dini.

BAB II

TINJAUAN PUSTAKA

Kajian pustaka memiliki peran penting dalam penelitian ilmiah karena berfungsi sebagai dasar teoretis yang memperkuat landasan konseptual dan metodologis dari penelitian yang dilakukan. Melalui kajian pustaka, peneliti dapat memahami sejauh mana penelitian-penelitian sebelumnya telah mengkaji topik yang serupa, menemukan kesenjangan penelitian, serta menentukan posisi penelitian yang sedang dilakukan dalam konteks keilmuan yang lebih luas. Dengan demikian, penyusunan landasan teori menjadi bagian krusial untuk menjamin bahwa penelitian memiliki arah yang jelas, relevan, dan berkontribusi terhadap pengembangan ilmu pengetahuan[5].

Dalam konteks penelitian ini, berbagai studi terdahulu telah membahas fenomena ketergantungan smartphone pada anak usia dini, baik dari perspektif psikologis, sosial, maupun pendidikan. Hasil penelitian menunjukkan bahwa penggunaan smartphone yang berlebihan pada anak dapat berdampak negatif terhadap perkembangan sosial, emosional, dan kognitif mereka. Selain itu, beberapa penelitian juga menyoroti bahwa pola penggunaan gawai yang tidak terkontrol dapat mengurangi kualitas interaksi anak dengan lingkungan sekitar maupun dengan orang tua. Temuan lainnya mengungkapkan bahwa anak yang terlalu sering menggunakan smartphone cenderung mengalami penurunan konsentrasi, kesulitan dalam mengatur emosi, serta gangguan pada pola tidur[6].

2.1 Penelitian Terdahulu

Berdasarkan hasil riview jurnal terdahulu, maka dapat perbandingan sebagaimana ditampilkan pada Tabel 2.1 di bawah ini:

Tabel 2. 1 Penelitian Terdahulu

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
1	**Rancang Bangun Sistem Prediksi Pola Sidik Jari Menggunakan Metode Backpropagation** *(Abdullah, Zikrullah, & Usman, 2023)* Jurnal Informatika dan Teknik Elektro Terapan (JITET), Vol. 11 No. 3.	Penelitian ini menghasilkan sebuah sistem prediksi pola sidik jari berbasis Backpropagation Neural Network (BPNN) yang dikembangkan melalui tahapan analisis kebutuhan, perancangan arsitektur jaringan, dan pengujian performa menggunakan MATLAB. Pada tahap awal, citra sidik jari dikumpulkan dan diproses melalui serangkaian *preprocessing* seperti konversi grayscale, normalisasi intensitas, *noise removal*, serta ekstraksi fitur garis lembut (ridge). Citra yang telah diproses kemudian dijadikan input bagi BPNN untuk proses pelatihan dan pengenalan pola.Arsitektur jaringan terdiri dari *input layer*, *hidden layer* yang diuji dalam beberapa konfigurasi neuron, serta *output layer* yang mewakili kelas pola sidik jari tertentu. Dari keseluruhan pengujian, diperoleh akurasi prediksi sebesar 89,74%, yang menunjukkan kemampuan model dalam mengenali pola biometrik secara otomatis. **Perbedaan:** Fokus penelitian ini adalah pengenalan pola citra biometrik, sedangkan penelitian yang saya lakukan saat ini berfokus pada analisis perilaku manusia, yaitu prediksi risiko ketergantungan smartphone pada anak usia dini menggunakan algoritma Random Forest. **Kelebihan:** Menunjukkan efektivitas jaringan saraf tiruan dalam pengenalan pola visual dengan hasil akurasi tinggi serta tahapan sistematis dalam implementasinya. **Kekurangan:** Model belum dibandingkan dengan algoritma lain dan terbatas pada jumlah data kecil, sehingga belum teruji untuk dataset besar dan beragam.

Tabel 2. 1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
2	**Analisis Komparatif Algoritma Berbasis Machine Learning untuk Memprediksi Tingkat Keparahan Cedera dalam Kecelakaan Lalu Lintas** Soumaya Amri, Mohammed Al Achhab, Mohamed Lazaar (2026))	Penelitian ini membandingkan algoritma CatBoost, HistGradientBoosting, Random Forest, dan SVM untuk memprediksi tingkat keparahan cedera akibat kecelakaan lalu lintas dengan menerapkan teknik class reduction dan SMOTE pada data yang tidak seimbang. **Perbedaan:** Penelitian tersebut menggunakan data kecelakaan lalu lintas dengan tujuan memprediksi tingkat keparahan cedera, sedangkan penelitian ini menggunakan data anak usia dini untuk memprediksi risiko ketergantungan smartphone. Meskipun objek penelitian berbeda, penelitian ini mengadopsi konsep penggabungan kelas (class reduction) pada algoritma Random Forest untuk mengatasi ketidakseimbangan data dan meningkatkan kinerja model. **Kelebihan:** Penelitian membandingkan empat algoritma (CatBoost, HistGradientBoosting, Random Forest, dan SVM) pada data yang tidak seimbang dengan menerapkan reduksi kelas dari empat kelas menjadi dua kelas (class reduction) serta SMOTE. Hasil menunjukkan bahwa Random Forest pada klasifikasi biner memperoleh performa terbaik dengan akurasi 83%, Macro F1 70%, dan F1 kelas parah 50%, sehingga reduksi kelas terbukti meningkatkan performa klasifikasi. **Kekurangan:** Penggabungan kelas menyebabkan hilangnya informasi tingkat keparahan yang lebih rinci karena empat kategori disederhanakan menjadi dua kategori. Selain itu, penelitian dilakukan pada data kecelakaan lalu lintas sehingga hasilnya belum tentu dapat digeneralisasi pada bidang lain.

Tabel 2. 1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
3	**Fenomena Kecanduan Gadget pada Anak Usia Pra Sekolah Akibat Pergeseran Nilai Budaya dan Faktor Kognitif Sosial** Mochamad Fauzan Putra Herdiana, Meilanny Budiarti Santoso (2024)	Hasil penelitian menunjukkan bahwa peningkatan penggunaan gadget pada anak pra-sekolah dipengaruhi oleh perubahan nilai dalam keluarga, terutama berkurangnya interaksi langsung antara orang tua dan anak akibat gaya hidup digital. Pergeseran nilai tersebut menyebabkan anak lebih banyak menerima *reinforcement* dari perangkat digital dibandingkan interaksi sosial nyata. Melalui perspektif kognitif sosial, penelitian ini menjelaskan bahwa perilaku anak terbentuk dari proses *observational learning*, di mana anak meniru kebiasaan penggunaan gadget yang dilihat dari orang tua atau lingkungan terdekat. **Perbedaan:**Penelitian Herdiana & Santoso (2024) bersifat kualitatif deskriptif dan teoritis**,** berfokus pada analisis sosial-budaya perilaku kecanduan gadget, sedangkan penelitian saya bersifat kuantitatif prediktif berbasis data, menggunakan algoritma *machine learning* Random Forest untuk memperkirakan tingkat risiko ketergantungan smartphone pada anak usia dini**.** **Kelebihan:** Kelebihan penelitian ini terletak pada kedalaman analisis teoritisnya**,** karena mampu mengaitkan dua konsep besar pergeseran nilai budaya dan teori kognitif sosial untuk menjelaskan secara menyeluruh penyebab kecanduan gadget pada anak. **Kekurangan:** Namun, penelitian ini memiliki keterbatasan karena bersifat deskriptif kualitatif tanpa dukungan data empiris atau kuantitatif, sehingga tidak menghasilkan model pengukuran tingkat ketergantungan gadget secara objektif. Selain itu, penelitian ini belum menjelaskan secara statistik faktor-faktor utama yang paling memengaruhi tingkat kecanduan, melainkan hanya mendeskripsikan secara konseptual hubungan antara perilaku orang tua dan anak.

Tabel 2. 1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
4	**Sistem Aplikasi Deteksi Tingkat Risiko Kehamilan pada AKI di Puskesmas Telaga Dewa** Reka Dwi Syaputra, Sari Widya, Deno Harmanto (2025)	Penelitian ini bertujuan untuk mengembangkan sistem aplikasi berbasis *machine learning* guna mendeteksi tingkat risiko kehamilan pada ibu hamil. Penelitian ini menggunakan kerangka kerja CRISP-DM (Cross Industry Standard Process for Data Mining) dengan menganalisis 488 data rekam medis ibu hamil yang terdiri atas 13 variabel klinis objektif, seperti tekanan darah, usia, kadar hemoglobin, dan riwayat abortus. Dua algoritma klasifikasi, yakni Decision Tree dan Random Forest, dibandingkan menggunakan teknik validasi silang *10-fold cross-validation* untuk memperoleh hasil yang optimal. **Perbedaan:**Perbedaan utamanya terletak pada objek dan konteks penelitian. Studi Syaputra dkk. (2025) berfokus pada deteksi risiko kehamilan dengan data klinis sebagai dasar klasifikasi, sedangkan penelitian ini berfokus pada prediksi risiko ketergantungan smartphone pada anak usia dini dengan menitikberatkan pada data perilaku dan sosial yang bersifat non-medis. **Kelebihan:** Penelitian ini memiliki kelebihan dalam penggunaan dua metode komparatif (Decision Tree dan Random Forest) yang dievaluasi secara sistematis melalui proses *hyperparameter tuning* dan validasi silang. Selain itu, penerapan kerangka CRISP-DM memastikan bahwa tahapan penelitian dilakukan secara terstruktur dan konsisten. **Kekurangan:** Namun demikian, penelitian ini memiliki keterbatasan karena hanya berfokus pada data klinis tanpa mempertimbangkan faktor sosial, psikologis, atau lingkungan yang turut memengaruhi risiko kehamilan. Selain itu, cakupan penelitian yang terbatas pada satu lokasi penelitian menyebabkan model yang dihasilkan belum dapat digeneralisasikan secara luas ke wilayah lain dengan karakteristik populasi yang berbeda.

Tabel 2.1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
5	**Pemanfaatan Media Pembelajaran Puzzle dalam Meningkatkan Kemampuan Kognitif Anak Usia Dini di TK Marina Permai** Hestia Alika K., Handriani J.H., Lisa Savatri Nataleni, Salini, Selvia Veronika, Wirastiani Binti Yusup (2025)	Penelitian menjelaskan bahwa anak usia dini merupakan anak yang berada pada rentang usia 1–6 tahun yang sedang mengalami masa pertumbuhan dan perkembangan yang sangat pesat (*golden age*), sehingga membutuhkan stimulasi yang tepat untuk mengoptimalkan perkembangan kognitifnya. Hasil penelitian menunjukkan bahwa penggunaan media pembelajaran puzzle efektif dalam meningkatkan kemampuan kognitif anak usia dini, seperti kemampuan mengenali bentuk dan warna, berpikir logis, memecahkan masalah, meningkatkan konsentrasi, daya ingat visual, ketekunan, serta rasa percaya diri. Guru menerapkan strategi pembelajaran yang disesuaikan dengan usia dan kemampuan anak sehingga proses belajar menjadi lebih interaktif dan menyenangkan. Selain meningkatkan aspek kognitif, media puzzle juga mendukung perkembangan sosial dan emosional anak. **Perbedaan:** Penelitian ini membahas media puzzle untuk meningkatkan kemampuan kognitif anak usia dini, sedangkan penelitian penulis berfokus pada prediksi risiko ketergantungan smartphone pada anak usia dini usia 1–6 tahun menggunakan algoritma Random Forest. **Kelebihan:** Menggunakan observasi, wawancara, dan dokumentasi sehingga hasil penelitian diperoleh secara mendalam serta disesuaikan dengan tahap perkembangan anak. **Kekurangan:** Penelitian hanya dilakukan pada satu TK dengan pendekatan kualitatif sehingga hasilnya belum dapat digeneralisasikan.

Tabel 2.1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
6	**Analisis Faktor Kenyamanan dan Ketergantungan Mahasiswa pada Smartphone yang Mengakibatkan Stres** Antonius Satria Hadi (2020)	Penelitian menjelaskan bahwa ketergantungan merupakan aktivitas yang dilakukan secara berulang-ulang sehingga menjadi kebiasaan dan dapat memberikan dampak positif maupun negatif. Dalam konteks smartphone, ketergantungan smartphone adalah perilaku menggunakan smartphone secara berulang dan cenderung berlebihan hingga menimbulkan dampak negatif. Hasil penelitian menunjukkan bahwa semakin tinggi tingkat kenyamanan dalam menggunakan smartphone, maka semakin tinggi pula tingkat ketergantungan smartphone. Selain itu, ketergantungan smartphone berpengaruh positif dan signifikan terhadap technostress, yaitu stres yang muncul akibat penggunaan teknologi secara berlebihan. **Perbedaan:** Penelitian ini menganalisis hubungan antara kenyamanan, ketergantungan smartphone, dan technostress pada mahasiswa. Sementara itu, penelitian penulis berfokus pada **prediksi risiko ketergantungan smartphone pada anak usia dini usia 1–6 tahun menggunakan algoritma Random Forest**, dengan variabel umur, durasi penggunaan smartphone, frekuensi penggunaan, usia pertama mengenal smartphone, tujuan penggunaan smartphone, dan pola pengawasan orang tua. **Kelebihan:** Menjelaskan konsep ketergantungan secara umum dan ketergantungan smartphone secara rinci serta didukung teori dan pengujian empiris menggunakan metode kuantitatif. **Kekurangan:** Penelitian hanya melibatkan mahasiswa sebagai responden dan berfokus pada pengaruh kenyamanan terhadap ketergantungan smartphone serta technostress, sehingga belum membahas faktor lain maupun kelompok anak usia dini.

Tabel 2.1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
7	**Penggunaan Gadget pada Anak Usia 4–5 Tahun dalam Keluarga dengan Pola Asuh Permisif** Sika Regiana, Lizza Suzanti, Rr. Deni Widjayatri (2024)	Penelitian yang dilakukan oleh Regiana, Suzanti, dan Widjayatri (2024) menunjukkan bahwa pola asuh permisif memiliki pengaruh signifikan terhadap tingginya intensitas penggunaan gadget pada anak usia 4–5 tahun. Anak diberikan keleluasaan tanpa adanya batasan durasi ataupun kontrol konten, sehingga membentuk kebiasaan penggunaan gadget yang cenderung berlebihan. Selain itu, penelitian ini menyoroti bagaimana kurangnya aturan dan pengawasan orang tua memicu munculnya perilaku pasif pada anak serta berkurangnya interesaksi sosial langsung. Temuan tersebut memperkuat pandangan bahwa pola asuh merupakan faktor dominan yang membentuk pola perilaku digital pada anak usia dini. **Perbedaan:**Perbedaan mendasar antara penelitian tersebut dengan penelitian yang penulis lakukan terletak pada pendekatannya. Jika penelitian Regiana dkk. bersifat deskriptif-kualitatif, maka penelitian penulis bersifat kuantitatif-prediktif dengan memanfaatkan algoritma *Random Forest* untuk memprediksi risiko ketergantungan smartphone pada anak usia dini berdasarkan variabel perilaku dan penggunaan teknologi digital. **Kelebihan:** Kelebihan dari penelitian ini terletak pada kemampuannya memberikan gambaran kontekstual yang mendalam mengenai hubungan antara pola asuh permisif dan perilaku anak terhadap penggunaan gadget, melalui pendekatan kualitatif yang mengungkap dinamika keluarga secara nyata. **Kekurangan:** Namun, penelitian ini memiliki keterbatasan karena hanya melibatkan satu subjek penelitian, sehingga hasilnya tidak dapat digeneralisasi. Selain itu, pendekatan kualitatif yang digunakan tidak memungkinkan pengukuran tingkat ketergantungan gadget secara objektif maupun analisis kuantitatif mengenai faktor penyebabnya.

Tabel 2.1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
8	**Prediksi Kesehatan Tidur dan Gaya Hidup Menggunakan Machine Learning** Mawadatul Maulidah, Nadiyah Hidayati (2024)	Penelitian Maulidah dan Hidayati (2024) bertujuan memprediksi gangguan tidur berdasarkan faktor gaya hidup menggunakan berbagai algoritma *machine learning* seperti Random Forest, SVM, KNN, dan Gradient Boosting. Dengan dataset *Sleep Health and Lifestyle* berisi 359 responden, hasil penelitian menunjukkan bahwa algoritma Gradient Boosting mencapai akurasi tertinggi sebesar 91%, sementara faktor durasi tidur, aktivitas fisik, stres, dan BMI menjadi variabel paling berpengaruh terhadap gangguan tidur. Temuan penelitian mengungkap bahwa durasi tidur, aktivitas fisik, tingkat stres, konsumsi alkohol, serta BMI memegang peranan besar dalam menentukan tingkat risiko gangguan tidur. Selain itu, peneliti menunjukkan bahwa model yang dikembangkan mampu mengelompokkan kategori kualitas tidur secara efektif, terutama pada kelas "Sehat" dan "Moderat". **Perbedaan:**Penelitian Maulidah dan Hidayati berfokus pada prediksi kesehatan tidur berdasarkan gaya hidup orang dewasa dengan algoritma Gradient Boosting, sedangkan penelitian penulis menitikberatkan pada prediksi risiko ketergantungan smartphone pada anak usia dini menggunakan algoritma Random Forest. **Kelebihan:** Kelebihan penelitian ini terletak pada penggunaan berbagai algoritma pembanding dengan proses *hyperparameter tuning* yang sistematis, serta pemanfaatan fitur gaya hidup yang beragam sehingga hasilnya dapat digunakan dalam pengembangan aplikasi kesehatan digital. **Kekurangan:** Namun demikian, penelitian ini masih memiliki keterbatasan pada jumlah data yang relatif kecil dan belum mencakup variasi populasi yang lebih luas. Selain itu, variabel psikologis seperti faktor sosial dan emosional belum dimasukkan dalam model, sehingga generalisasi hasilnya masih terbatas.

Tabel 2.1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
9	**Implementasi Algoritma K-Nearest Neighbor Untuk Prediksi Padi Nasional: Studi Kasus Pada Data Produksi Padi Indonesia** Muhammad Ady Nugroho, Pratiwi Cahyaningtiyas, Rizqi Aufa Eka Prathama, Rengga Arga Deva, Arif Setiawan (2025)	Penelitian oleh **Nugroho dkk. (2025)** berjudul *"Implementasi Algoritma K-Nearest Neighbor untuk Prediksi Produksi Padi Nasional"* bertujuan memprediksi produksi padi di Indonesia menggunakan data historis BPS 2018–2023. Model K-Nearest Neighbor (K = 2) dievaluasi dengan MSE dan R^2 Score, menghasilkan nilai R^2 sebesar 0,88 yang menunjukkan akurasi tinggi. Hasil penelitian menunjukkan bahwa model KNN menghasilkan nilai R^2 sebesar 0,88, yang menandakan kemampuan model dalam menangkap pola tren produksi secara cukup akurat. Model kemudian digunakan untuk memprediksi produksi padi tahun 2024–2025, dan hasil prediksi memperlihatkan tren yang cenderung stabil, konsisten dengan pola historis yang diamati. **Perbedaan:** Nugroho dkk. fokus pada prediksi produksi padi nasional menggunakan algoritma KNN, sedangkan penelitian penulis berfokus pada prediksi risiko ketergantungan smartphone pada anak usia dini menggunakan algoritma Random Forest. **Kelebihan:** Kelebihan penelitian ini terletak pada penggunaan data resmi dari BPS yang kredibel serta penerapan algoritma KNN yang sederhana namun mampu memberikan hasil prediksi yang akurat. Selain itu, visualisasi hasil prediksi disajikan dengan baik, sehingga memudahkan pemangku kebijakan memahami tren yang terjadi. **Kekurangan:** Namun, kelemahan penelitian ini adalah terbatasnya variabel yang digunakan, karena model hanya mempertimbangkan faktor waktu dan produksi tanpa melibatkan variabel eksternal seperti cuaca, luas lahan, atau kebijakan pemerintah yang berpotensi memengaruhi hasil prediksi. Nilai *K* yang kecil dan ukuran dataset yang terbatas juga dapat menurunkan kemampuan model dalam melakukan generalisasi terhadap data baru.

Tabel 2.1 Penelitian Terdahulu Lanjutan

NO	Judul, Penulis, dan Tahun	Hasil Penelitian (Kelebihan, Kekurangan, Perbedaan)
10	**Prediksi Kemampuan Akademik Mhasiswa dengan Metode Support Vector Machine** Bangkit Indarmawan Nugroho, Nugroho Adhi Santoso, Aang Alim Murtopo (2023)	Penelitian oleh **Nugroho, Santoso, dan Murtopo (2023)** berjudul *"Prediksi Kemampuan Akademik Mahasiswa dengan Metode Support Vector Machine"* bertujuan membangun model prediksi kemampuan akademik berdasarkan faktor demografi, *multiple intelligence*, dan *learning style* menggunakan algoritma SVM kernel polynomial orde-2. Dengan data 204 mahasiswa AMIK YMI Tegal, model mencapai akurasi 92,5%, dan faktor logical-mathematical intelligence (MI 02) terbukti paling berpengaruh dengan akurasi 97,5%. **Perbedaan:**Perbedaan terletak pada objek dan metode: penelitian ini memprediksi kemampuan akademik mahasiswa menggunakan SVM, sedangkan penelitian penulis memprediksi risiko ketergantungan smartphone pada anak usia dini menggunakan Random Forest yang lebih unggul dalam menangani variabel kompleks dan menghasilkan interpretasi faktor yang terukur. **Kelebihan:** Kelebihan penelitian ini terletak pada penerapan algoritma SVM dengan kernel polynomial yang efektif dalam mengklasifikasikan data nonlinier serta menghasilkan tingkat akurasi tinggi. Selain itu, penelitian ini berhasil menggabungkan aspek psikologis dan gaya belajar dalam analisis prediksi akademik, sehingga memberikan pandangan yang lebih komprehensif terhadap faktor-faktor penentu prestasi mahasiswa. **Kekurangan:** Namun, penelitian ini masih memiliki keterbatasan pada cakupan data yang sempit, karena hanya melibatkan satu institusi (AMIK YMI Tegal), serta belum membandingkan performa SVM dengan algoritma lain seperti Random Forest atau Neural Network.

Berdasarkan Tabel 2.1, dapat disimpulkan bahwa berbagai penelitian sebelumnya telah menerapkan metode *machine learning* pada beragam bidang

seperti kesehatan, pendidikan, sosial, dan teknologi. Secara umum, seluruh penelitian tersebut menunjukkan bahwa algoritma *machine learning*, terutama *Random Forest*, memiliki kemampuan prediktif yang tinggi dan akurat dalam mengklasifikasikan data yang kompleks.

Penelitian pertama di lakukan oleh [7] berhasil membangun sistem prediksi pola sidik jari menggunakan *Backpropagation Neural Network (BPNN)* dengan akurasi 89,74%. Hasil ini memperlihatkan potensi besar jaringan saraf tiruan dalam mengenali pola visual, meskipun masih terbatas pada konteks biometrik.

Penelitian selanjutnya dilakukan oleh [8] menunjukkan efektivitas beberapa algoritma machine learning seperti Random Forest, Logistic Regression, Naïve Bayes, SVM, dan Neural Networks dalam memprediksi risiko stunting. Meskipun mampu membandingkan performa model dengan baik, penelitian ini belum mengkaji faktor-faktor paling berpengaruh terhadap hasil prediksi. Fokus penelitian yang hanya pada perbandingan akurasi tanpa analisis fitur penting menjadikannya berbeda dengan penelitian saya, yang menitikberatkan pada eksplorasi faktor dominan penyebab ketergantungan.

Berbeda dengan penelitian-penelitian berbasis data numerik dan medis tersebut, studi yang dilakukan [9] dan yang dilakukan oleh [10] menyoroti fenomena kecanduan gadget dari perspektif sosial dan psikologis melalui pendekatan kualitatif. Kedua penelitian ini memberikan pemahaman teoritis mendalam tentang penyebab dan dampak kecanduan gadget, namun belum mengembangkan model kuantitatif yang dapat mengukur tingkat ketergantungan secara objektif.

Dari sisi metodologis, penelitian yang dilakukan [11] serta penelitian yang dilakukan [12] menampilkan variasi penerapan algoritma seperti *Gradient Boosting*, *K-Nearest Neighbor*, dan *Support Vector Machine* di bidang kesehatan, pertanian, dan pendidikan. Hasilnya menegaskan bahwa *machine learning* memiliki fleksibilitas tinggi dalam berbagai konteks data, meskipun sebagian besar penelitian masih terbatas pada data deterministik dan tidak melibatkan variabel perilaku manusia.

Dengan demikian, penelitian ini hadir untuk mengisi celah penelitian yang belum disentuh oleh penelitian terdahulu. Fokus penelitian ini adalah mengembangkan model prediksi risiko ketergantungan smartphone pada anak usia dini menggunakan algoritma Random Forest, yang tidak hanya menghasilkan klasifikasi risiko (rendah, sedang, tinggi), tetapi juga berpotensi mampu menunjukkan faktor-faktor paling berpengaruh terhadap munculnya ketergantungan tersebut.

Pendekatan ini diharapkan dapat menjadi kontribusi baru dalam pemanfaatan *machine learning* pada bidang sosial dan psikologis, khususnya dalam isu perkembangan perilaku anak di era digital.

2.2 Ketergantungan Smartphone

Ketergantungan smartphone merupakan fenomena perilaku yang semakin banyak ditemukan pada berbagai kelompok usia, termasuk anak usia dini. Secara konseptual, ketergantungan smartphone dapat dipahami sebagai kondisi di mana individu mengalami kesulitan dalam mengontrol penggunaan perangkat tersebut hingga menimbulkan dampak negatif terhadap aktivitas sosial, emosional, dan kognitif. [13] menjelaskan bahwa ketergantungan smartphone termasuk dalam

kategori *behavioral addiction* atau kecanduan non-substansi, yang ditandai oleh dorongan kompulsif untuk terus menggunakan perangkat, kehilangan kendali, serta munculnya rasa gelisah atau stres ketika akses terhadap smartphone dibatasi. Pada anak usia dini, gejala ini tampak dalam bentuk perilaku rewel, mudah marah, atau enggan beraktivitas tanpa kehadiran smartphone.

Dalam kajian teoretis, ketergantungan smartphone dapat dijelaskan melalui beberapa model dan teori perilaku. Teori *Uses and Gratifications* yang dikemukakan oleh [14] menyatakan bahwa individu menggunakan media untuk memenuhi kebutuhan tertentu seperti hiburan, informasi, dan interaksi sosial. Dalam konteks anak usia dini, penggunaan smartphone sering kali menjadi sarana pemenuhan kebutuhan akan hiburan dan rasa nyaman, baik sebagai bentuk kesenangan pribadi maupun karena dorongan lingkungan keluarga. Sementara itu, teori *Addictive Behavior* oleh [15] menekankan bahwa perilaku adiktif muncul akibat adanya penguatan positif (*positive reinforcement*), yaitu perasaan senang dan kepuasan yang dirasakan setelah menggunakan media, yang kemudian mendorong pengulangan perilaku tersebut secara terus-menerus.

Berdasarkan kedua teori tersebut, ketergantungan smartphone pada anak usia dini dapat dipahami sebagai bentuk interaksi kompleks antara kebutuhan psikologis, lingkungan sosial, dan proses pembiasaan yang diperkuat oleh rangsangan digital. Kondisi ini menunjukkan pentingnya pendekatan ilmiah yang lebih objektif dalam memahami dan mengukur tingkat ketergantungan smartphone. Oleh karena itu, penelitian ini berupaya mengembangkan model prediksi berbasis algoritma *Random Forest* untuk mengidentifikasi risiko ketergantungan smartphone pada anak usia dini secara kuantitatif dan terukur,

sehingga dapat menjadi dasar dalam upaya pencegahan perilaku adiktif di era digital.

2.3 Anak Usia Dini

Anak usia dini merupakan kelompok individu yang berada pada tahap awal perkembangan kehidupan manusia, yang memiliki karakteristik khas baik dari aspek fisik, kognitif, sosial, maupun emosional. Berdasarkan Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 137 Tahun 2014, anak usia dini adalah anak yang berada pada rentang usia 0 hingga 6 tahun. Sedangkan menurut World Health Organization (WHO), anak usia dini mencakup rentang usia 0 sampai 8 tahun, yaitu masa ketika pertumbuhan otak dan pembentukan kepribadian berkembang secara pesat. Pada periode ini, anak berada dalam masa *golden age* yang sangat menentukan kualitas perkembangan di masa depan.

Perkembangan teknologi digital saat ini telah menjangkau hampir semua lapisan masyarakat, termasuk anak usia dini. Smartphone, sebagai salah satu bentuk teknologi digital yang paling populer, kini telah menjadi bagian dari kehidupan sehari-hari anak-anak. Menurut Kominfo (2023), sekitar 65% anak usia 5–8 tahun di Indonesia telah terpapar penggunaan smartphone, baik untuk menonton video, bermain gim, maupun mengakses aplikasi pembelajaran. Penggunaan smartphone pada anak umumnya diawali dari keinginan orang tua untuk memberikan hiburan atau sarana edukatif yang mudah diakses. Namun, seiring waktu, kebiasaan ini sering berkembang menjadi pola penggunaan yang berlebihan, terutama ketika anak mulai mengasosiasikan smartphone sebagai sumber utama kesenangan dan kenyamanan.

Dalam konteks ini, peran orang tua menjadi sangat penting dalam mengatur, mengawasi, dan membimbing penggunaan smartphone pada anak usia dini. Menurut teori *Parental Mediation* yang dikemukakan oleh [16] terdapat tiga bentuk utama pengawasan orang tua terhadap penggunaan media digital, yaitu *restrictive mediation* (pembatasan durasi dan jenis konten), *active mediation* (pendampingan dan diskusi tentang konten yang diakses), serta *co-viewing* (menggunakan media bersama anak). Pola asuh yang terlalu permisif atau kurang pengawasan dapat meningkatkan risiko anak mengalami ketergantungan terhadap smartphone.

Dengan demikian, perilaku penggunaan smartphone pada anak usia dini merupakan hasil interaksi antara faktor perkembangan individu dan pola pengasuhan keluarga. Jika penggunaan smartphone dilakukan secara seimbang, maka perangkat ini dapat menjadi sarana edukatif yang bermanfaat. Namun, tanpa pengawasan yang tepat, smartphone justru berpotensi menimbulkan perilaku adiktif yang berdampak negatif pada perkembangan sosial, emosional, dan kognitif anak. Oleh karena itu, pemahaman terhadap karakteristik anak usia dini dan pola penggunaan teknologi digital sangat penting dalam pengembangan model prediksi risiko ketergantungan smartphone, sebagaimana diupayakan dalam penelitian ini melalui penerapan algoritma *Random Forest* untuk menganalisis faktor-faktor yang memengaruhi tingkat ketergantungan anak secara objektif dan terukur.

2.4 Machine Learning

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence/AI) telah membawa kemajuan signifikan dalam berbagai bidang, salah satunya

melalui cabang utamanya yaitu *machine learning*. *Machine learning* merupakan metode analisis data yang memungkinkan sistem komputer untuk belajar dari data secara otomatis tanpa harus diprogram secara eksplisit. Menurut [17] *machine learning* adalah suatu proses di mana komputer memiliki kemampuan untuk meningkatkan kinerjanya berdasarkan pengalaman yang diperoleh dari data. Dengan demikian, sistem dapat mengenali pola, membuat keputusan, dan melakukan prediksi terhadap data baru berdasarkan pembelajaran dari data sebelumnya.

Dalam konteks penelitian ilmiah modern, *machine learning* sering dimanfaatkan untuk membangun model prediksi, klasifikasi, maupun deteksi pola kompleks yang sulit dianalisis secara manual. Model ini bekerja dengan mengidentifikasi hubungan antarvariabel dalam data untuk menghasilkan suatu keputusan yang optimal. Menurut [18] terdapat tiga kategori utama algoritma *machine learning*, yaitu supervised learning, unsupervised learning, dan ensemble learning.

Pertama, supervised learning adalah pendekatan pembelajaran yang menggunakan data berlabel, di mana model dilatih berdasarkan pasangan data input dan output yang sudah diketahui. Contoh algoritma yang termasuk dalam kategori ini antara lain *Decision Tree*, *Support Vector Machine (SVM)*, *Naïve Bayes*, *K-Nearest Neighbor (KNN)*, dan *Random Forest*. Jenis algoritma ini sering digunakan dalam masalah klasifikasi dan regresi, termasuk prediksi perilaku manusia atau kecenderungan sosial.

Kedua, unsupervised learning bekerja tanpa data berlabel, sehingga model berfungsi untuk menemukan pola tersembunyi atau mengelompokkan data

berdasarkan kemiripan karakteristik. Contoh algoritma yang umum digunakan adalah *K-Means Clustering*, *Hierarchical Clustering*, dan *Principal Component Analysis (PCA)*. Pendekatan ini biasanya digunakan untuk eksplorasi data dan segmentasi perilaku dalam populasi tertentu.

Ketiga, ensemble learning merupakan metode yang menggabungkan beberapa model *machine learning* untuk meningkatkan akurasi dan stabilitas hasil prediksi. Salah satu algoritma ensemble yang paling populer adalah *Random Forest*, yang bekerja dengan menggabungkan banyak *Decision Tree* dan menggunakan hasil voting mayoritas untuk menentukan prediksi akhir. Pendekatan ini terbukti lebih tahan terhadap *overfitting* dan mampu menangani data yang memiliki banyak variabel atau hubungan nonlinier.

Dalam bidang prediksi perilaku sosial dan psikologis, penerapan *machine learning* semakin banyak digunakan karena kemampuannya mengolah data yang kompleks dan beragam, seperti faktor demografis, perilaku digital, interaksi sosial, hingga variabel psikologis. Melalui analisis data perilaku dan pola penggunaan teknologi, *machine learning* dapat mengidentifikasi variabel-variabel penting yang memengaruhi kecenderungan seseorang terhadap suatu perilaku tertentu.

Dalam konteks penelitian ini, penerapan *machine learning* berfokus pada pengembangan model prediksi untuk mengidentifikasi risiko ketergantungan smartphone pada anak usia dini. Penggunaan algoritma *Random Forest* dipilih karena kemampuannya dalam menangani berbagai jenis variabel (numerik maupun kategorikal), serta menghasilkan interpretasi terhadap *feature importance* yang dapat menjelaskan faktor-faktor dominan penyebab ketergantungan. Dengan

demikian, penggunaan *machine learning* tidak hanya memberikan pendekatan komputasional dalam memahami perilaku digital anak, tetapi juga berkontribusi terhadap upaya pencegahan dini dan pembentukan kebijakan edukatif yang berbasis data.

2.5 Algoritma Random Forest

Algoritma Random Forest merupakan salah satu metode *ensemble learning* yang banyak digunakan dalam *machine learning* untuk menyelesaikan permasalahan klasifikasi dan regresi. Menurut [19] *Random Forest* adalah kumpulan dari beberapa *decision tree* yang dibangun secara acak dari subset data dan fitur yang berbeda, kemudian hasil prediksi dari seluruh pohon tersebut digabungkan melalui proses voting (untuk klasifikasi) atau rata-rata (untuk regresi). Konsep utama dari algoritma ini adalah menggabungkan banyak model sederhana (*weak learners*) menjadi satu model yang kuat (*strong learner*) guna meningkatkan akurasi dan kestabilan hasil prediksi.

Secara prinsip dasar, setiap *decision tree* dalam Random Forest dilatih menggunakan teknik bootstrap sampling, yaitu pengambilan sampel data secara acak dengan pengembalian (*sampling with replacement*). Setiap pohon hanya dilatih dengan sebagian data dan sebagian fitur yang dipilih secara acak. Hal ini menyebabkan setiap pohon memiliki karakteristik yang unik dan tidak saling bergantung. Ketika proses pelatihan selesai, seluruh pohon keputusan tersebut akan menghasilkan prediksi terhadap data uji, dan hasil akhir diperoleh berdasarkan mayoritas suara (majority voting) untuk kasus klasifikasi atau nilai rata-rata (mean prediction) untuk kasus regresi. Pendekatan ini membuat Random Forest lebih tahan terhadap *overfitting* dibandingkan model tunggal seperti

Decision Tree, karena variasi antar-pohon membantu menyeimbangkan kesalahan prediksi.

Menurut [20] keunggulan utama Random Forest terletak pada kemampuannya dalam menangani dataset yang besar, kompleks, dan mengandung variabel dengan tipe berbeda (baik numerik maupun kategorikal). Selain itu, algoritma ini memiliki kelebihan dalam mengukur tingkat kepentingan setiap fitur (*feature importance*) yang berkontribusi terhadap hasil prediksi. Informasi tersebut berguna bagi peneliti untuk memahami variabel mana yang paling berpengaruh terhadap suatu fenomena. Random Forest juga memiliki performa yang stabil meskipun terdapat data yang hilang (*missing values*) atau terdapat *outlier* dalam dataset.

Namun demikian, Random Forest juga memiliki beberapa kelemahan. Pertama, karena terdiri dari banyak *decision tree*, model ini relatif membutuhkan waktu komputasi dan memori yang lebih besar dibandingkan algoritma sederhana seperti *Logistic Regression* atau *Naïve Bayes*. Kedua, meskipun interpretasi *feature importance* dapat dilakukan, model secara keseluruhan tergolong *black box model*, sehingga sulit untuk menjelaskan secara rinci bagaimana keputusan akhir terbentuk di setiap pohon. Selain itu, jika parameter jumlah pohon (*n_estimators*) terlalu banyak, proses pelatihan dapat menjadi lambat dan tidak efisien.

Penerapan algoritma Random Forest telah banyak digunakan dalam berbagai bidang penelitian, baik di ranah medis, ekonomi, pendidikan, maupun sosial. Dalam bidang kesehatan, misalnya, Random Forest digunakan untuk

memprediksi risiko penyakit jantung dan diabetes, berdasarkan data klinis pasien [21], [22].

Dalam konteks penelitian ini, algoritma Random Forest diterapkan untuk membangun model prediksi risiko ketergantungan smartphone pada anak usia dini. Melalui kemampuannya dalam menangani data kompleks dan melakukan analisis *feature importance*, Random Forest diharapkan mampu mengidentifikasi faktor-faktor utama yang berkontribusi terhadap munculnya perilaku ketergantungan, seperti durasi penggunaan smartphone, jenis konten yang diakses, serta tingkat pengawasan orang tua. Dengan demikian, hasil prediksi yang diperoleh dapat memberikan dasar ilmiah yang kuat dalam upaya pencegahan dini terhadap ketergantungan digital pada anak usia dini.

2.6 Variabel Dependen

Dalam penelitian ini, variabel dependen yang digunakan adalah tingkat risiko ketergantungan smartphone pada anak usia dini. Variabel ini menjadi fokus utama yang ingin diprediksi menggunakan algoritma Random Forest. Dengan kata lain, variabel dependen merupakan hasil atau keluaran dari proses analisis yang menunjukkan sejauh mana seorang anak mengalami ketergantungan terhadap penggunaan smartphone, yang dapat dikategorikan ke dalam tingkat rendah dan tinggi.

Hubungan variabel dependen dengan penelitian ini sangat erat, karena tujuan utama dari penelitian adalah memprediksi nilai atau tingkat risiko ketergantungan tersebut berdasarkan berbagai faktor yang memengaruhinya. Faktor-faktor tersebut, yang disebut sebagai variabel independen, meliputi antara lain umur, durasi penggunaan smartphone, tujuan penggunaan smartphone, pola

pengawasan orang tua, frekuensi penggunaan smartphone, dan usia pertama kali mengenal smartphone.

Melalui penerapan algoritma Random Forest, penelitian ini berupaya untuk mengidentifikasi pola hubungan antara variabel independen dengan variabel dependen, sehingga sistem mampu menghasilkan prediksi yang akurat mengenai tingkat risiko ketergantungan smartphone. Dengan demikian, variabel dependen berperan penting sebagai tujuan akhir dari proses analisis dan prediksi dalam penelitian ini, sekaligus menjadi dasar dalam memberikan rekomendasi pencegahan atau intervensi bagi anak usia dini yang berpotensi mengalami ketergantungan terhadap smartphone.

Variabel ini dikategorikan dalam dua tingkatan risiko, yaitu:

1. Rendah: anak menggunakan smartphone secara terbatas dan di bawah pengawasan orang tua.
2. Tinggi: anak menunjukkan perilaku penggunaan berlebihan hingga mengganggu aktivitas sehari-hari, interaksi sosial, dan emosi.

Data tingkat risiko diperoleh melalui kuesioner terstruktur berdasarkan indikator perilaku penggunaan smartphone dan hasil klasifikasi model *machine learning* menggunakan algoritma Random Forest.

2.7 Variabel Independen

Variabel independent adalah variabel yang mempengaruhi atau menjadi faktor penyebab perubahan pada variabel dependen. Dalam penelitian berjudul "Prediksi Risiko Ketergantungan Smartphone pada Anak Usia Dini Menggunakan Algoritma Random Forest", variabel independen merupakan faktor-faktor yang

diduga berkontribusi terhadap tingkat risiko ketergantungan smartphone pada anak usia dini.

variabel independen dalam penelitian ini antara lain:

1. Umur
2. Durasi Penggunaan Smartphone
3. Tujuan Penggunaan Smartphone
4. Pola Pengawasan Orang Tua
5. Frekuensi penggunaan smartphone
6. Usia pertama kali mengenal smartphone

Hubungannya dengan penelitian ini adalah bahwa variabel-variabel independen tersebut menjadi input atau masukan bagi algoritma Random Forest untuk melakukan proses analisis dan klasifikasi. Melalui data dari variabel-variabel ini, algoritma akan mencari pola dan hubungan yang dapat digunakan untuk memprediksi nilai dari variabel dependen, yaitu tingkat risiko ketergantungan smartphone.

Dengan demikian, keenam variabel independen tersebut akan digunakan sebagai masukan (*input features*) dalam model *machine learning* berbasis algoritma Random Forest untuk memprediksi variabel dependen berupa tingkat risiko ketergantungan smartphone pada anak usia dini. Pemilihan variabel ini didasarkan pada kajian teoretis dan hasil penelitian terdahulu yang menunjukkan hubungan erat antara faktor usia, durasi penggunaan, pola asuh, serta aspek emosional dengan kecenderungan ketergantungan digital pada anak.

2.8 Rangkuman

Berdasarkan hasil telaah terhadap sepuluh jurnal relevan, dapat disimpulkan bahwa penerapan algoritma *machine learning* telah banyak digunakan dalam berbagai bidang, terutama kesehatan, pendidikan, agrikultur, dan lingkungan. Namun demikian, penggunaan algoritma prediktif untuk menganalisis perilaku digital anak usia dini masih sangat terbatas. Sebagian besar penelitian terdahulu berfokus pada data numerik, medis, atau fisik yang bersifat deterministik, bukan pada aspek sosial dan psikologis yang kompleks.

Dari hasil telaah tersebut dapat diidentifikasi bahwa kesenjangan penelitian terletak pada belum adanya penerapan algoritma *machine learning*, khususnya *Random Forest*, untuk menganalisis dan memprediksi perilaku ketergantungan digital anak usia dini**.** Mayoritas penelitian terdahulu hanya menyoroti kecanduan gadget secara deskriptif atau menerapkan model prediksi pada data medis dan fisik, tanpa mengintegrasikan faktor sosial, psikologis, dan perilaku anak.

Kebaruan penelitian ini terletak pada aspek konteks, variabel, pendekatan model, dan kontribusi praktisnya. Penelitian ini mengalihkan penerapan *machine learning* dari bidang medis, agrikultur, dan pendidikan menuju domain perilaku digital anak usia dini**,** yang masih jarang dikaji secara kuantitatif. Selain itu, penelitian ini mengintegrasikan variabel sosial dan psikologis yang lebih spesifik, seperti durasi penggunaan smartphone, jenis aplikasi, pola pengawasan orang tua, serta faktor lingkungan keluarga dan perilaku emosional anak — variabel-variabel yang sebelumnya belum banyak dimasukkan dalam model prediktif berbasis data.

Dari sisi metodologis, penelitian ini menghadirkan pendekatan baru melalui penerapan algoritma ensemble learning Random Forest untuk menganalisis risiko perilaku manusia. Algoritma ini memiliki kemampuan lebih baik dalam menangani variabel yang kompleks dan saling berkorelasi dibandingkan metode tunggal seperti *SVM*, *KNN*, atau *ANN* yang umum digunakan dalam penelitian terdahulu. Dengan demikian, penelitian ini tidak hanya berfokus pada akurasi prediksi tingkat risiko, tetapi juga memberikan kontribusi ilmiah dan praktis dalam upaya pencegahan dini ketergantungan digital anak melalui model prediksi yang objektif, adaptif, dan berbasis data perilaku nyata.

Selain itu, variabel-variabel penelitian dalam studi ini memiliki landasan empiris yang kuat berdasarkan penelitian terdahulu. Variabel durasi penggunaan smartphone**,** jenis aplikasi, dan pengawasan orang tua didukung oleh studi perilaku keluarga [10] sedangkan variabel perilaku emosional anak berakar dari kajian psikologis yang membahas kecanduan gadget [9].

Dengan demikian, seluruh variabel penelitian memiliki dasar teoritis dan empiris yang relevan, mendukung pembangunan model prediksi risiko ketergantungan smartphone pada anak usia dini menggunakan algoritma Random Forest yang lebih komprehensif dan aplikatif di era digital.

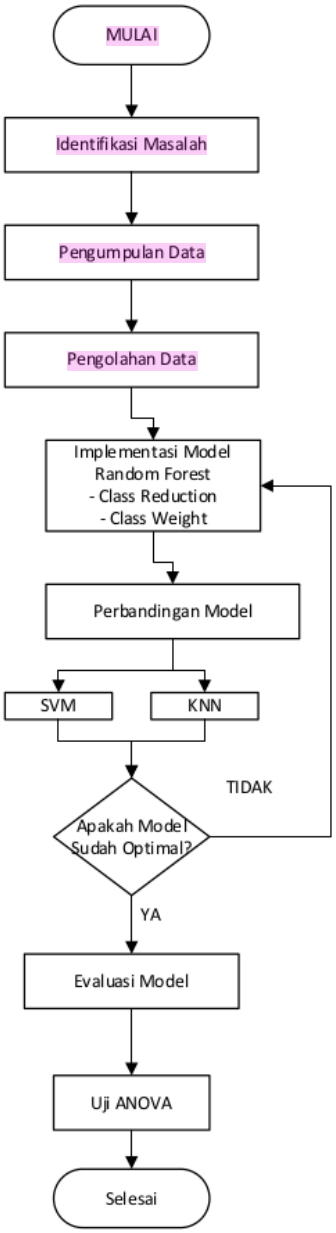
BAB III

METODE PENELITIAN

Dalam bab metode penelitian ini, berisi langkah-langkah yang akan digunakan dalam memprediksi risiko ketergantungan smartphone pada anak usia dini, yang menjelaskan langkah-langkah mulai dari mengidentifikasi masalah hingga selesai. Pada penelitian ini metode pengumpulan data menggunakan metode kuantitatif, model yang digunakan yaitu Algoritma Random Forest sebagai alat untuk prediksi.

3.1 Kerangka Penelitian

Kerangka penelitian menggambarkan alur berpikir serta tahapan kegiatan yang dilakukan peneliti untuk mencapai tujuan penelitian. Setiap langkah disusun secara sistematis agar proses pengumpulan, pengolahan, dan analisis data dapat berlangsung secara terarah dan menghasilkan model prediksi yang akurat dan dapat dipertanggungjawabkan. Dalam penelitian ini, tahapan yang dilakukan mencakup pengumpulan data, preprocessing dan pengolahan data, serta implementasi algoritma Random Forest sebagai model utama. Pada tahap evaluasi, model dinilai menggunakan metrik seperti accuracy, precision, recall, F1-score, serta confusion matrix untuk mengukur performanya secara menyeluruh. Selanjutnya, hasil prediksi dianalisis untuk mengidentifikasi variabel atau fitur yang paling berpengaruh terhadap risiko ketergantungan smartphone pada anak usia dini. Keseluruhan proses ini memberikan landasan yang kuat bagi penarikan kesimpulan dan rekomendasi penelitian.



Gambar 3.1 Alur Penelitian

3.1.1 Penjelasan Alur Penelitian

1. Mulai

Penelitian diawali dengan menentukan fokus dan arah penelitian yang akan dilakukan.

2. Identifikasi Masalah

Pada tahap ini peneliti mengidentifikasi masalah utama yang akan diteliti, yaitu bagaimana memprediksi atau mengklasifikasikan suatu fenomena menggunakan algoritma machine learning dengan *Random Forest* sebagai algoritma utama.

Identifikasi masalah menjadi dasar untuk menentukan kebutuhan data, metode, serta evaluasi yang tepat.

3. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan melalui dua sumber, yaitu data primer dan data sekunder, yang digunakan untuk membangun serta mengevaluasi model prediksi risiko ketergantungan smartphone pada anak usia dini.

1. Data Primer

Data primer diperoleh melalui penyebaran kuesioner kepada orang tua yang memiliki anak usia dini (1–6 tahun). Kuesioner tersebut disusun berdasarkan variabel penelitian, antara lain: durasi penggunaan smartphone, jenis aplikasi yang digunakan, pengawasan orang tua, perilaku agresif, kecemasan dan depresif, serta perilaku menarik diri.

Kuesioner menggunakan skala Likert dan skala penilaian adaptasi *Child Behavior Checklist (CBCL)* untuk mengukur karakteristik perilaku anak secara terstandar. Data primer ini digunakan untuk memperoleh gambaran langsung mengenai perilaku penggunaan smartphone dan kondisi emosional anak.

2. Data Sekunder

Data sekunder diperoleh dari dataset publik di platform Kaggle serta hasil penelitian terdahulu yang relevan. Dataset sekunder digunakan sebagai pembanding dan pelengkap dalam proses pelatihan model, terutama dalam tahap eksplorasi pola, pengujian fitur, dan validasi model. Penggunaan data sekunder membantu memperkaya variasi data, menambah jumlah sampel, serta meningkatkan generalisasi model prediksi.

4. Pengolahan Data (Preprocessing)

Data yang telah terkumpul dibersihkan dan dipersiapkan melalui beberapa proses seperti:

1) Pembersihan data (cleaning)
2) Encoding
3) Normalisasi/standarisasi

Tahap ini memastikan data yang masuk ke model adalah data yang berkualitas.

5. Implementasi Model Random Forest

Random Forest digunakan sebagai algoritma utama dalam penelitian.

Prosesnya meliputi:

1) Melakukan train–test split
2) Melatih model Random Forest pada data training
3) Melakukan prediksi dengan data testing
4) Menghasilkan nilai akurasi atau metrik performa lainnya

Hasil performa Random Forest menjadi acuan awal sebelum dilakukan perbandingan.

6. Perbandingan Model

Setelah Random Forest diimplementasikan, penelitian dilanjutkan dengan membangun dua model lain, yaitu:

1) Support Vector Machine (SVM)
2) K-Nearest Neighbors (KNN)

Tujuan tahap ini adalah melihat apakah Random Forest merupakan model yang paling optimal dibandingkan model lain.

7. Evaluasi Model

Semua model (Random Forest, SVM, KNN) dievaluasi menggunakan metrik evaluasi yang sama, yaitu:

Accuracy (Akurasi): Persentase prediksi yang benar terhadap total data.

Rumus:

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP + TN}}{\boldsymbol{TP + TN + FP + FN}}$$

1) *Precision*: Proporsi hasil positif yang benar di antara semua hasil yang diprediksi positif.

   Rumus:

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP + FP}}$$

2) *Recall* (Sensitivitas): Kemampuan model dalam mendeteksi seluruh kasus positif yang benar.

   Rumus:

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP + FN}}$$

3) *F1-Score*: Rata-rata harmonis dari *precision* dan *recall* untuk mengukur keseimbangan kinerja model.

Rumus:

$$F1 - Score = 2 * \frac{precision * Recall}{Precission * Recall}$$

4) *Confusion Matrix*: Tabel yang menggambarkan jumlah prediksi benar dan salah untuk setiap kelas.

Tabel 3.1 Confusion Matrix

	Prediksi Positif	**Prediksi Negatif**
Aktual Positif	TP	FN
Aktual Negatif	FP	TN

Tabel Confusion Matrix menggambarkan hubungan antara kelas aktual dan prediksi pada model klasifikasi. Nilai True Positive (TP) dan True Negative (TN) menunjukkan jumlah prediksi yang benar untuk masing-masing kelas positif dan negatif. Sementara itu, False Positive (FP) dan False Negative (FN) merepresentasikan kesalahan prediksi, yaitu ketika model salah mengklasifikasikan data negatif sebagai positif maupun sebaliknya. Tabel ini digunakan sebagai dasar dalam perhitungan metrik evaluasi seperti akurasi, precision, recall, dan F1-score.

8. Uji ANOVA

Setelah mendapatkan nilai performa dari ketiga model, dilakukan Uji ANOVA untuk membuktikan secara statistik akurasi antar model.

Langkah-Langkah Uji ANOVA:

1) Mengumpulkan nilai akurasi dari setiap algoritma sebanyak k pengujian (misalnya 5-fold / 10-fold).
2) Menghitung rata-rata dan varians dari masing-masing algoritma.
3) Menghitung nilai F hitung (F-statistic) menggunakan rumus:

$$F = \frac{SSB|(k-1)}{SSW|(N-k}$$

Dengan:

SSB (Sum of Squares Between): variasi antar algoritma

SSW (Sum of Squares Within): variasi dalam tiap algoritma

k = jumlah algoritma (4 algoritma)

N = total seluruh sampel evaluasi (misalnya 4 algoritma × 10 kali pengujian)

Menentukan nilai p-value dari F-statistik.

Menarik kesimpulan:

Jika p-value < 0.05 → *Terdapat perbedaan signifikan performa antar algoritma.*

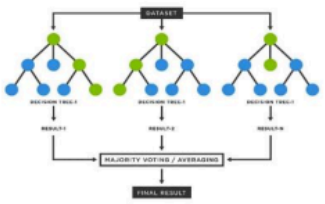
Jika p-value ≥ 0.05 → *Tidak terdapat perbedaan signifikan.*

9. Selesai

Penelitian diakhiri dengan hasil prediksi dan penarikan kesimpulan dari seluruh proses, yaitu model mana yang paling optimal dan bagaimana hasil uji statistik mendukung temuan tersebut.

3.2 Model Random Forest

Proses pemodelan Random Forest dalam penelitian ini dimulai dari tahap persiapan hingga tahap evaluasi akhir. Setiap langkah dilakukan secara sistematis agar model yang dihasilkan mampu memberikan prediksi yang akurat dan dapat dipertanggungjawabkan secara ilmiah. Berikut merupakan gambaran umum dari alur kerja Random Forest yang di tunjukan pada gambar 3.2.



Gambar 3.2 Random Forest

Sumber: https://medium.com

1. Dataset sebagai input utama

   Proses dimulai dari dataset yang telah disiapkan. Dataset ini menjadi sumber informasi yang akan dipelajari oleh seluruh pohon keputusan dalam model Random Forest.

2. Pembentukan beberapa pohon keputusan

   Dataset kemudian dibagi menjadi beberapa subset dengan teknik *bootstrap sampling*

3. Masing-masing pohon menghasilkan prediksi awal

   Setelah dilatih, setiap pohon memberikan hasil prediksi mandiri.

   Output dari setiap pohon merupakan representasi keputusan berdasarkan struktur cabangnya masing-masing.

4. Proses majority voting / voting terbanyak

   Semua prediksi dari pohon-pohon tersebut digabungkan menggunakan metode *majority voting*.

   Pada tahap ini, kelas yang paling sering muncul dalam prediksi pohon-pohon akan dianggap sebagai keputusan kolektif.

5. Penentuan prediksi akhir (*Final Result*)

   Hasil voting menjadi prediksi akhir *Random Forest*.

   Karena berasal dari gabungan banyak pohon, keputusan akhir menjadi lebih stabil dan memiliki tingkat akurasi yang lebih baik dibandingkan hanya menggunakan satu pohon Keputusan.

BAB IV
HASIL PENELITIAN

4.1 Data Penelitian

Pada penelitian ini, data yang digunakan terdiri dari data primer dan data sekunder. Data primer diperoleh melalui penyebaran kuesioner kepada responden yang merupakan orang tua dari anak usia dini, sedangkan data sekunder diperoleh dari sumber pendukung seperti jurnal dan dataset publik.

Jumlah responden yang digunakan dalam penelitian ini sebanyak 347 responden. Data yang dikumpulkan mencakup beberapa variabel yang berhubungan dengan penggunaan smartphone pada anak usia dini. Variabel tersebut terdiri dari satu variabel dependen yaitu tingkat risiko ketergantungan smartphone, serta beberapa variabel independen seperti umur, durasi penggunaan smartphone, tujuan penggunaan smartphone, pola pengawasan orang tua, frekuensi penggunaan smartphone, usia pertama kali mengenal smartphone. Berdasarkan hasil pengumpulan data, mayoritas anak menggunakan smartphone dengan durasi lebih dari 2 jam per hari, serta jenis aplikasi yang dominan digunakan adalah aplikasi hiburan seperti video dan game. Hal ini menunjukkan adanya potensi risiko ketergantungan yang cukup tinggi pada anak usia dini.

4.2 Analisis Risiko Ketergantungan Smartphone Berdasarkan Umur

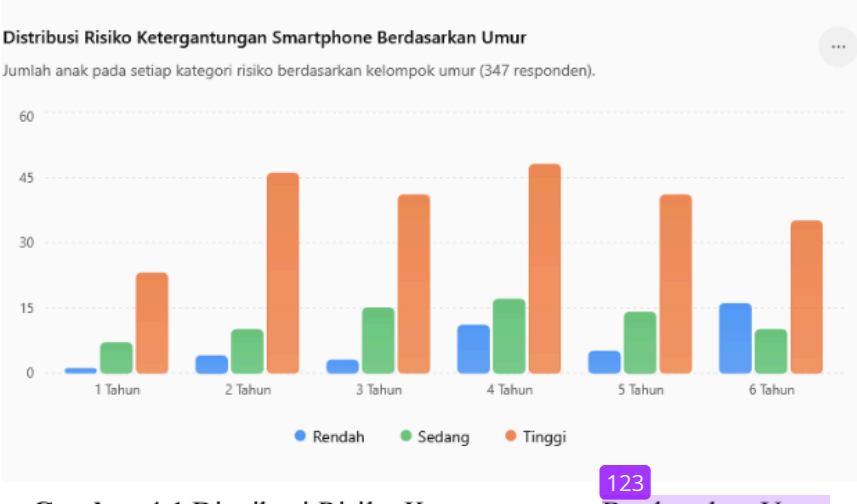
Variabel umur merupakan salah satu variabel independen yang digunakan dalam penelitian ini untuk memprediksi risiko ketergantungan smartphone pada anak usia dini. Berdasarkan hasil analisis terhadap 347 responden, diperoleh distribusi tingkat risiko ketergantungan smartphone pada setiap kelompok umur sebagaimana ditunjukkan pada Tabel 4.1

Tabel 4.1 Distribusi Risiko Ketergantungan Berdasarkan Umur

Umur	Rendah	Sedang	Tinggi	Total
1 tahun	1	7	23	31
2 tahun	4	10	46	60
3 tahun	3	15	41	59
4 tahun	11	17	48	76
5 tahun	5	14	41	60
6 tahun	16	10	35	61
Total	**40**	**73**	**234**	**347**

Tabel 4.2 Tingkat Risiko Ketergantungan Berdasarkan Rentang Usia Anak

Rentang Usia	Hasil Penelitian	Interpretasi
1–2 tahun	Risiko tinggi mulai terlihat, dengan 23 anak usia 1 tahun dan 46 anak usia 2 tahun berada pada kategori risiko tinggi.	Anak mulai terpapar smartphone sejak usia dini sehingga diperlukan pendampingan orang tua.
3–4 tahun	Jumlah risiko tinggi meningkat menjadi 41 anak pada usia 3 tahun dan mencapai puncaknya pada usia 4 tahun sebanyak 48 anak.	Rentang usia 3–4 tahun merupakan periode yang paling rentan berdasarkan jumlah kasus pada penelitian ini.
5–6 tahun	Risiko tinggi masih ditemukan (41 anak usia 5 tahun dan 35 anak usia 6 tahun), namun jumlah anak dengan risiko rendah juga meningkat.	Anak masih berpotensi mengalami ketergantungan, tetapi sebagian telah menunjukkan penggunaan yang lebih terkontrol melalui pengawasan orang tua.



Gambar 4.1 Distribusi Risiko Ketergantungan Berdasarkan Umur

Berdasarkan Tabel 4.2, dapat diketahui bahwa risiko ketergantungan smartphone mulai terlihat sejak usia 1–2 tahun dan meningkat pada rentang usia 3–4 tahun. Jumlah anak dengan kategori risiko tinggi mencapai nilai tertinggi pada usia 4 tahun, yaitu sebanyak 48 anak. Setelah memasuki usia 5–6 tahun, kategori risiko tinggi masih mendominasi, namun jumlah anak dengan kategori risiko rendah mulai meningkat. Hal ini menunjukkan bahwa rentang usia 3–4 tahun merupakan periode yang paling rentan terhadap risiko ketergantungan smartphone berdasarkan hasil penelitian ini, sehingga diperlukan pengawasan orang tua yang lebih optimal pada rentang usia tersebut.

4.3 Preprocessing Data

Tahap preprocessing dilakukan untuk memastikan kualitas data sebelum digunakan dalam proses pemodelan. Tahapan yang dilakukan meliputi data cleaning, transformasi data, normalisasi data, dan pembagian data.

Pada tahap data cleaning, dilakukan pengecekan terhadap data yang tidak lengkap (*missing value*) dan duplikasi data. Hasil pengecekan menunjukkan bahwa seluruh data telah lengkap dan tidak ditemukan data duplikat, sehingga data dapat langsung digunakan pada tahap berikutnya.Selanjutnya dilakukan proses encoding untuk mengubah data kategorikal menjadi bentuk numerik agar dapat diproses oleh algoritma machine learning. Proses encoding dilakukan menggunakan metode *Label Encoding*, seperti kelas Rendah = 0, Sedang = 1, dan Tinggi = 2.

Nama orang Tua/Wali Anak	1. Berapa usia anak Anda saat ini?	2. Rata-rata berapa jam anak Anda menggu	3. Seberapa sering anak Anda menggunakan
Aira hafsah	2 Tahun	Kurang dari 1 jam	1 kali
Widia	4 Tahun	2–3 jam	4–5 kali
Fitri Kafriyana	3 Tahun	Kurang dari 1 jam	1 kali
Fajar Dwi P	6 Tahun	1–2 jam	2–3 kali
Neni Hermita	5 Tahun	2–3 jam	4–5 kali
ANGGIT UTAMI	3 Tahun	1–2 jam	2–3 kali
NANI YULIA BS	6 Tahun	Lebih dari 3 jam	Lebih dari 5 kali
Hagi Ahmad F	4 Tahun	1–2 jam	2–3 kali

Gambar 4.1 Data Asli

4. Pada usia berapa anak Anda mulai mengg	5. Aktivitas apa saja yang paling sering dilak	6. Apakah Anda memiliki aturan khusus terk	7. Dalam skala 1 (tidak pernah) sampai 5 (selalu), s
2 Tahun	Menonton video (YouTube, dll)	Ya, selalu diterapkan	5
3 Tahun	Menonton video (YouTube, dll), Bermain gam	Ya, kadang-kadang diterapkan	3
2 Tahun	Belajar (aplikasi edukasi)	Ya, selalu diterapkan	5
4 Tahun	Menonton video (YouTube, dll), Belajar (aplika	Ya, selalu diterapkan	5
3 Tahun	Menonton video (YouTube, dll), Bermain gam	Ya, kadang-kadang diterapkan	4
2 Tahun	Menonton video (YouTube, dll), Belajar (aplika	Ya, selalu diterapkan	4
2 Tahun	Menonton video (YouTube, dll), Bermain gam	Tidak memiliki aturan khusus	2
3 Tahun	Menonton video (YouTube, dll), Belajar (aplika	Ya, selalu diterapkan	4

Gambar 4.2 Data Asli

8. Alasan utama apa yang mendasari Anda m	9. Anak mudah marah ketika penggunaan sm	10. Anak menunjukkan perilaku agresif (mena	11. Anak lebih memilih bermain dengan smar
Sebagai sarana hiburan	2	2	1
Agar anak tidak bosan/rewel	3	3	3
Untuk edukasi atau pembelajaran	5	5	5
Untuk edukasi atau pembelajaran	5	5	5
Sebagai sarana hiburan, Agar anak tidak bos	5	4	4
Untuk edukasi atau pembelajaran	4	5	4
Sebagai sarana hiburan, Agar anak tidak bos	2	1	2
Untuk edukasi atau pembelajaran	5	5	5

Gambar 4.3 Data Asli

12. Ai				
1	1	1	1	2
4	3	4	3	3
5	4	5	4	4
4	4	5	5	4
2	5	4	5	4
4	4	4	4	4
2	2	2	2	1
5	5	5	4	5

Gambar 4.4 Data Asli

Umur	Durasi	Frekuensi	Usia Awal	Tujuan	Pengawasan Orang Tua	Tingkat Risiko Ketergantungan
2	1	1	2	0,5	4	Rendah
4	3	3	3	1,5	2,5	Sedang
3	1	1	2	5	4	Tinggi
6	2	2	4	3	4	Tinggi
5	3	3	3	3,5	3	Sedang
3	2	2	2	3	3,5	Sedang
6	4	4	2	3,5	1,5	Rendah
4	2	2	3	3	3,5	Tinggi
4	2	2	3	3,5	4	Tinggi
5	2	2	4	3,5	3,5	Tinggi
5	3	3	3	3,5	3	Tinggi
3	1	1	2	3	4	Tinggi
4	2	2	2	3,5	3,5	Tinggi
3	1	1	2	3,5	4	Tinggi
2	3	3	1	3	3	Tinggi
5	1	1	4	3	4	Tinggi
6	4	4	4	3	1	Rendah

Gambar 4. 5 Data Setelah Encoding

4.4 Pembagian Data Training dan Data Testing

Sebelum dilakukan proses pelatihan model, data penelitian terlebih dahulu dibagi menjadi data training dan data testing menggunakan metode split data dengan perbandingan 80:20. Dari total 347 data, sebanyak 277 data digunakan sebagai data training dan 70 data digunakan sebagai data testing.

Tabel 4. 1 Distribusi Kelas Keseluruhan

Kelas	Jumlah Data
Rendah	40
Sedang	73
Tinggi	234
Total	**347**

Berdasarkan distribusi data, diketahui bahwa jumlah data pada setiap kelas masih belum seimbang. Kelas risiko tinggi memiliki jumlah data paling banyak yaitu 234 data, sedangkan kelas risiko sedang sebanyak 73 data dan kelas risiko rendah sebanyak 40 data. Ketidakseimbangan data ini dapat mempengaruhi performa model dalam melakukan klasifikasi karena model cenderung lebih mudah mengenali kelas mayoritas dibandingkan kelas minoritas.

Tabel 4. 2 Pembagian Data Training dan Data Testing

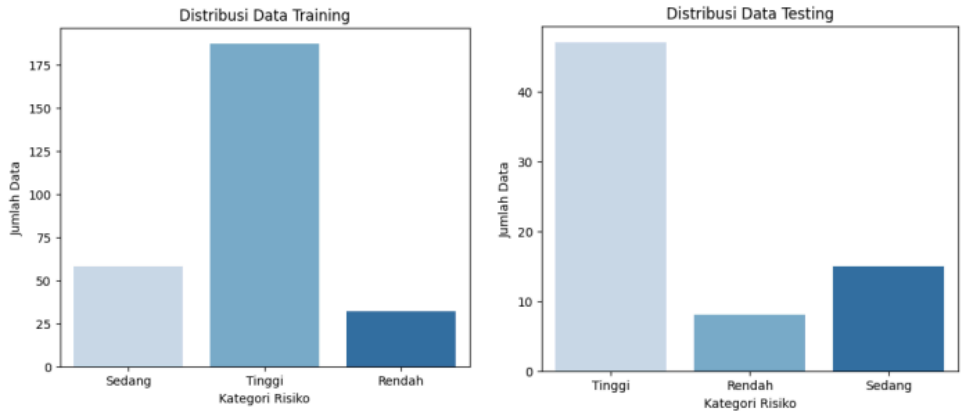
Jenis Data	Jumlah
Training	277
Testing	70

Tabel 4.3 Distribusi Data Training

Kelas	Jumlah Data
Rendah	32
Sedang	58
Tinggi	187
Total	**277**

Tabel 4. 4 Distribusi Data Testing

Kelas	Jumlah Data
Rendah	8
Sedang	15
Tinggi	47
Total	**70**



Gambar 4.6 Distribusi Data Training dan Data Testing

Berdasarkan hasil pembagian data, diperoleh sebanyak 277 data training dan 70 data testing. Pada data training, kelas risiko tinggi memiliki jumlah data paling banyak yaitu 187 data, sedangkan kelas risiko sedang sebanyak 58 data dan kelas risiko rendah sebanyak 32 data.

Sementara itu, pada data testing terdapat 47 data untuk kelas risiko tinggi, 15 data untuk kelas risiko sedang, dan 8 data untuk kelas risiko rendah. Hasil distribusi tersebut menunjukkan bahwa jumlah data pada setiap kelas masih tidak seimbang, di mana kelas risiko tinggi mendominasi dibandingkan kelas lainnya.

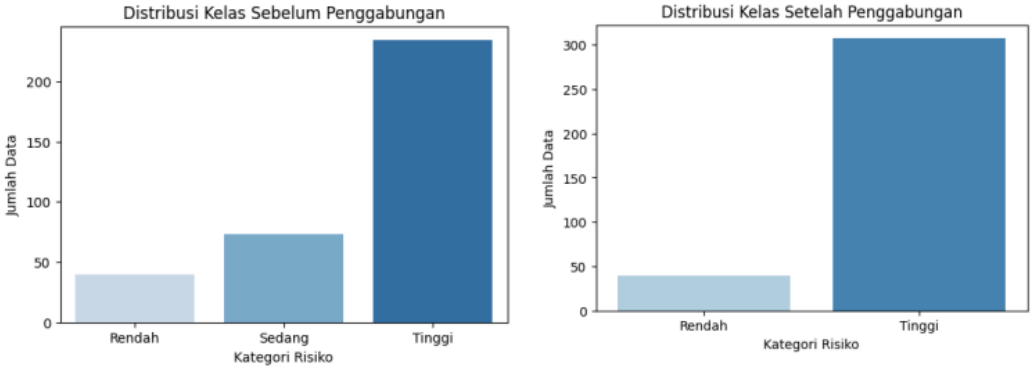
4.5 Pelatihan Model

4.5.1 Model Random Forest

1. Dataset sebagai Input Utama

Proses klasifikasi dimulai dengan menggunakan dataset hasil preprocessing yang terdiri dari 347 data. Dataset tersebut memiliki enam

atribut yaitu Umur, Durasi, Frekuensi, Usia Awal, Tujuan, dan Pengawasan Orang Tua, dengan variabel target berupa Tingkat Risiko Ketergantungan Smartphone. Sebelum proses pelatihan model dilakukan, diterapkan teknik penggabungan kelas (*class reduction*) dengan menggabungkan kelas Sedang ke dalam kelas Tinggi. Teknik ini digunakan untuk mengurangi ketidakseimbangan data serta meningkatkan kemampuan model dalam mengenali pola klasifikasi.



Gambar 4.7 Distribusi Kelas Sebelum dan Setelah Penggabungan

Berdasarkan hasil penggabungan kelas, jumlah data pada kategori Tinggi meningkat menjadi 307 data, sedangkan kategori Rendah tetap sebanyak 40 data.

2. Pembentukan Beberapa Pohon Keputusan

   Setelah dataset siap digunakan, data dibagi menjadi data training dan data testing dengan perbandingan 80:20. Sebanyak 277 data digunakan sebagai data training dan 70 data digunakan sebagai data testing.

   Model Random Forest kemudian dibangun menggunakan parameter sebagai berikut:

Tabel 4.5 Parameter

Parameter	Nilai
n_estimators	500
max_depth	20
min_samples_split	2
min_samples_leaf	1
max_features	sqrt
class_weight	balanced_subsample

Pada tahap ini, Random Forest membentuk 500 pohon keputusan menggunakan teknik bootstrap sampling. Setiap pohon dilatih menggunakan subset data yang berbeda sehingga menghasilkan model yang lebih stabil dan mampu mengurangi overfitting.

3. Masing-Masing Pohon Menghasilkan Prediksi Awal

Setelah proses pelatihan selesai, setiap pohon keputusan melakukan prediksi terhadap data testing secara independen. Setiap pohon menghasilkan keputusan berdasarkan pola yang dipelajari dari data training.

Sebagai contoh, suatu data uji dapat menghasilkan prediksi yang berbeda dari beberapa pohon keputusan.

Tabel 4.6 Contoh Pohon Keputusan

Pohon Keputusan	Prediksi
Pohon 1	Tinggi
Pohon 2	Tinggi
Pohon 3	Rendah
Pohon 4	Tinggi
Pohon 5	Tinggi

Prediksi dari masing-masing pohon kemudian dikumpulkan untuk menentukan keputusan akhir model.

4. Proses Majority Voting (Voting Terbanyak)

Misalkan terdapat satu data testing yang akan diprediksi oleh model Random Forest. Dari total 500 pohon keputusan yang terbentuk, masing-masing pohon memberikan hasil prediksi sebagai berikut:

Tabel 4.7 Contoh Prediksi Pohon Keputusan

Pohon Keputusan	**Prediksi**
Pohon 1	Tinggi
Pohon 2	Tinggi
Pohon 3	Rendah
Pohon 4	Tinggi
Pohon 5	Tinggi

Berdasarkan contoh hasil prediksi tersebut, diperoleh jumlah suara sebagai berikut:

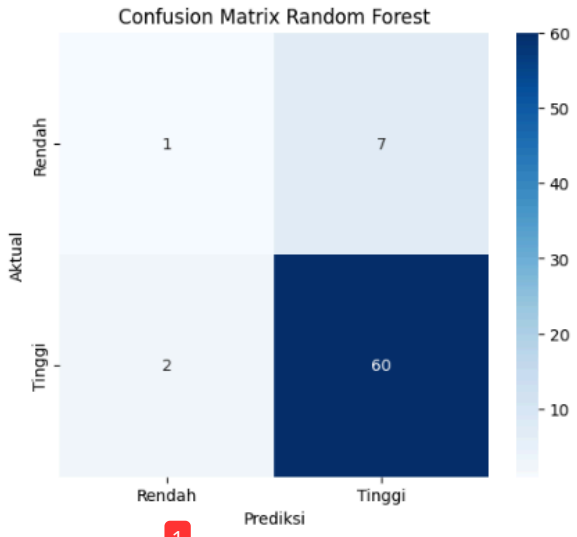
Tabel 4.8 Jumlah Prediksi

Kelas	**Jumlah Suara**
Rendah	1
Tinggi	4

Berdasarkan ilustrasi pada Tabel 4.7 dan 4.8 tersebut, setiap pohon keputusan menghasilkan prediksi secara independen. Seluruh prediksi kemudian digabungkan menggunakan metode majority voting. Karena kelas Tinggi memperoleh suara terbanyak (4 suara), maka prediksi akhir yang dihasilkan model adalah Tinggi. Metode ini membantu meningkatkan stabilitas dan akurasi klasifikasi Random Forest.

5. Penentuan Prediksi Akhir (Final Result)

Hasil voting digunakan sebagai prediksi akhir model Random Forest. Berdasarkan pengujian menggunakan 70 data testing, diperoleh confusion matrix sebagai berikut.



Gambar 4.8 Confusion Matrix Random Forest

Berdasarkan confusion matrix tersebut, sebanyak 1 data kategori Rendah dan 60 data kategori Tinggi berhasil diklasifikasikan dengan benar. Sementara itu, terdapat 7 data kategori Rendah yang salah diprediksi sebagai Tinggi dan 2 data kategori Tinggi yang salah diprediksi sebagai Rendah.

Tabel 4. 9 Classification Report *Random Forest*

Kelas	**Precision**	**Recall**	**F1-Score**	**Support**
Rendah	0,33	0,12	0,18	8
Tinggi	0,90	0,97	0,93	62
Accuracy	-	-	**0,87**	**70**
Macro Average	0,61	0,55	0,56	70
Weighted Average	0,83	0,87	0,84	70

Untuk memastikan nilai evaluasi yang dihasilkan oleh model Random Forest, dilakukan perhitungan manual berdasarkan confusion matrix. Perhitungan

ini meliputi precision, recall, F1-score, accuracy, macro average, dan weighted average.

Tabel 4. 10 Nilai TP, FP dan FN

Kelas	TP	FP	FN
Rendah	1	2	7
Tinggi	60	7	2

1. Perhitungan Precision

   Rumus: $Precision = \frac{TP}{TP+FP}$

a. Precision kelas Rendah

$$= \frac{1}{1+2}$$

$$= \frac{1}{3} = 0{,}33$$

b. Precision kelas Tinggi

$$= \frac{60}{60+7}$$

$$= \frac{60}{67} = 0{,}90$$

2. Perhitungan Recall

   Rumus: $Recall = \frac{TP}{TP+FN}$

   a. Recall kelas Rendah

$$= \frac{1}{1+7}$$

$$= \frac{1}{8} = 0{,}125 \approx 0{,}12$$

   b. Recall kelas Tinggi

$$= \frac{60}{60+2}$$

$$= \frac{60}{62} = 0{,}968 \approx 0{,}97$$

3. Perhitungan F1-Score

Rumus: $F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$

a. F1-Score kelas Rendah

$$= 2 \times \frac{0{,}33 \times 0{,}12}{0{,}33 + 0{,}12}$$

$$= 2 \times \frac{0{,}0396}{0{,}45}$$

$$= \frac{0{,}0792}{0{,}45} = 0{,}176 \approx 0{,}18$$

b. F1-Score kelas Tinggi

$$= 2 \times \frac{0{,}90 \times 0{,}97}{0{,}90 + 0{,}97}$$

$$= 2 \times \frac{0{,}873}{1{,}87}$$

$$= \frac{1{,}746}{1{,}87} = 0{,}934 \approx 0{,}93$$

4. Perhitungan Accuracy

Rumus: $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

Prediksi benar:

Rendah = 1

Tinggi = 60

Total benar = 61

Jumlah data = 70

$$Accuracy = \frac{61}{70}$$

$$= 0{,}8714$$

$$= 87{,}14\% \approx 87\%$$

5. Perhitungan Macro Averege

Precision Macro: $Precision_{Macro} = \frac{\sum Precision_i}{n}$

$$= \frac{0{,}33 + 0{,}90}{2}$$

$$= \frac{1{,}23}{2}$$

$$= 0{,}615 \approx 0{,}61$$

Recall Macro: $Recall_{Macro} = \frac{\sum Recall_i}{n}$

$$= \frac{0{,}12 + 0{,}97}{2}$$

$$= \frac{1{,}09}{2}$$

$$= 0{,}545 \approx 0{,}55$$

F1-Score Mcaro: $F1_{Macro} = \frac{\sum F1_i}{n}$

$$= \frac{0{,}18 + 0{,}93}{2}$$

$$= \frac{1{,}11}{2}$$

$$= 0{,}555 \approx 0{,}56$$

6. Perhitungan Weighted Averege

Perhitungan weighted average mempertimbangkan jumlah data (support) pada masing-masing kelas.

Precision Weighted: $Precision_{Weighted} = \frac{\sum(Precision_i \times Support_i)}{\sum Support_i}$

$$= \frac{(0{,}33 \times 8) + (0{,}90 \times 62)}{70}$$

$$= \frac{2,64 + 55,80}{70}$$

$$= \frac{58,44}{70}$$

$$= 0,835 \approx 0,83$$

Recall Weighted: $Recall_{Weighted} = \frac{\sum(Recall_i \times Support_i)}{\sum Support_i}$

$$= \frac{(0,12 \times 8) + (0,97 \times 62)}{70}$$

$$= \frac{0,96 + 60,14}{70}$$

$$= \frac{61,10}{70}$$

$$= 0,873 \approx 0,87$$

F1-Score Weighted: $F1_{Weighted} = \frac{\sum(F1_i \times Support_i)}{\sum Support_i}$

$$= \frac{(0,18 \times 8) + (0,93 \times 62)}{70}$$

$$= \frac{1,44 + 57,66}{70}$$

$$= \frac{59,10}{70}$$

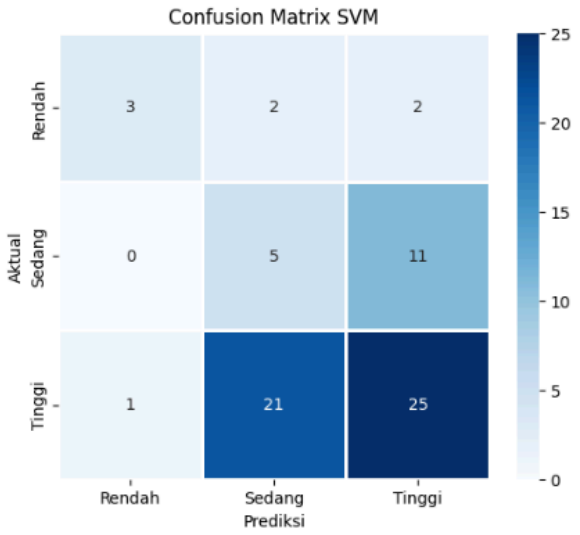
$$= 0,844 \approx 0,84$$

12 Berdasarkan hasil perhitungan tersebut, model Random Forest memperoleh accuracy sebesar 87,14%, yang menunjukkan bahwa model mampu mengklasifikasikan 61 dari 70 data uji dengan benar. 13 Pada kelas Tinggi, model memiliki precision 0,90, recall 0,97, dan F1-score 0,93, sehingga menunjukkan kemampuan yang sangat baik dalam mengenali data berisiko tinggi. Sementara itu, pada kelas Rendah, diperoleh 13 precision 0,33, recall 0,12, dan F1-score 0,18, yang menunjukkan bahwa model masih mengalami kesulitan dalam

mengidentifikasi kelas minoritas karena jumlah datanya jauh lebih sedikit dibandingkan kelas Tinggi. Nilai macro average sebesar 0,61 (precision), 0,55 (recall), dan 0,56 (F1-score) menggambarkan rata-rata performa pada kedua kelas tanpa mempertimbangkan jumlah data, sedangkan weighted average sebesar 0,83 (precision), 0,87 (recall), dan 0,84 (F1-score) menunjukkan bahwa performa keseluruhan model didominasi oleh kelas Tinggi yang memiliki jumlah data lebih banyak.

4.5.2 Model Support Vector Machine (SVM)

Model Support Vector Machine digunakan sebagai algoritma pembanding dalam penelitian ini.



Gambar 4.9 Confusion Matrix SVM

Berdasarkan hasil confusion matrix pada algoritma SVM, diketahui bahwa model masih belum mampu melakukan klasifikasi dengan optimal. Pada kelas Rendah, model berhasil memprediksi 3 data dengan benar, namun terdapat 4 data yang salah klasifikasi. Pada kelas Sedang, hanya 5 data yang benar diprediksi, sedangkan 11 data salah diklasifikasikan menjadi kelas Tinggi. Sementara pada

kelas Tinggi, model berhasil memprediksi 25 data dengan benar, tetapi masih terdapat 22 data yang salah klasifikasi.

Tabel 4. 11 Classification Report Model SVM

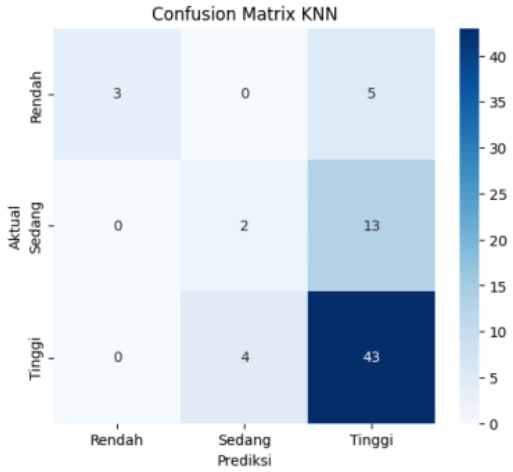
Kelas	Precision	Recall	F1-Score	Support
Rendah	0,75	0,43	0,55	7
Sedang	0,18	0,31	0,23	16
Tinggi	0,66	0,53	0,59	47
Accuracy			**0,47**	70
Macro Average	0,53	0,42	0,45	70
Weighted Average	0,56	0,47	0,50	70

Berdasarkan hasil classification report pada Tabel 4.10, model Support Vector Machine (SVM) memperoleh accuracy sebesar 47% pada 70 data pengujian. Pada kelas Rendah, model menghasilkan precision sebesar 0,75, recall sebesar 0,43, dan F1-score sebesar 0,55, yang menunjukkan bahwa prediksi untuk kelas Rendah cukup tepat, namun masih terdapat beberapa data yang belum berhasil dikenali. Pada kelas Sedang, diperoleh precision sebesar 0,18, recall sebesar 0,31, dan F1-score sebesar 0,23, yang menunjukkan bahwa model masih mengalami kesulitan dalam mengklasifikasikan kelas Sedang. Sementara itu, pada kelas Tinggi, model memperoleh precision sebesar 0,66, recall sebesar 0,53, dan F1-score sebesar 0,59, sehingga menunjukkan kemampuan yang cukup baik dalam mengenali kelas Tinggi meskipun masih terdapat sejumlah kesalahan klasifikasi.

Secara keseluruhan, nilai macro average sebesar 0,53 untuk precision, 0,42 untuk recall, dan 0,45 untuk F1-score serta weighted average masing-masing sebesar 0,56, 0,47, dan 0,50 menunjukkan bahwa performa model SVM masih tergolong rendah dalam mengklasifikasikan tingkat risiko ketergantungan smartphone pada anak usia dini dibandingkan model Random Forest.

4.5.3 Model K-Nearest Neighbor (KNN)

Berdasarkan hasil pengujian metode K-Nearest Neighbor (KNN), dilakukan percobaan menggunakan nilai K dari 1 hingga 5 untuk menentukan parameter terbaik pada model. Hasil pengujian menunjukkan bahwa nilai akurasi tertinggi diperoleh pada K = 5 dengan nilai akurasi sebesar 69%, sehingga nilai tersebut dipilih sebagai parameter optimal dalam proses klasifikasi.



Gambar 4.10 Confusion Matrix KNN

Berdasarkan hasil confusion matrix KNN, pada kelas Rendah sebanyak 3 dari 8 data berhasil diklasifikasikan dengan benar, sedangkan 5 data salah diprediksi menjadi kelas Tinggi. Pada kelas Sedang, hanya 2 dari 15 data yang berhasil diprediksi dengan benar, sementara 13 data salah diklasifikasikan menjadi kelas Tinggi. Sementara itu, pada kelas Tinggi, sebanyak 43 dari 47 data berhasil diprediksi dengan benar, sedangkan 4 data salah diprediksi menjadi kelas Sedang. Secara keseluruhan, hasil confusion matrix menunjukkan bahwa model KNN lebih baik dalam mengenali kelas Tinggi, namun masih kurang optimal dalam membedakan kelas Rendah dan Sedang.

Tabel 4. 12 Classification Report Model KNN

Kelas	Precision	Recall	F1-Score	Support
Rendah	1,00	0,38	0,55	8
Sedang	0,33	0,13	0,19	15
Tinggi	0,70	0,91	0,80	47
Accuracy			**0,69**	70
Macro Average	0,68	0,47	0,51	70
Weighted Average	0,66	0,69	0,64	70

Jika ditinjau berdasarkan classification report, pada kelas Rendah diperoleh nilai precision sebesar 1,00, recall sebesar 0,38, dan f1-score sebesar 0,55. Nilai recall yang rendah menunjukkan bahwa model masih mengalami kesulitan dalam mengenali seluruh data pada kategori Rendah, meskipun prediksi yang dihasilkan sudah sangat tepat.

Pada kelas Sedang, diperoleh nilai precision sebesar 0,33, recall sebesar 0,13, dan f1-score sebesar 0,19. Nilai tersebut menunjukkan bahwa model belum mampu mengidentifikasi data pada kelas Sedang dengan baik, karena sebagian besar data masih salah diklasifikasikan ke kelas lain.

Sementara itu, pada kelas Tinggi, model menunjukkan performa yang cukup baik dengan nilai precision sebesar 0,70, recall sebesar 0,91, dan f1-score sebesar 0,80, sehingga sebagian besar data pada kategori Tinggi berhasil diprediksi dengan benar oleh model.

4.6 Perbandingan Performa Model

Berdasarkan hasil pengujian, diperoleh perbandingan performa antar model sebagai berikut:

Tabel 4. 13 Perbandingan Performa Model

Model	Accuracy	Precision (Macro Avg)	Recall (Macro Avg)	F1-Score (Macro Avg)
Random Forest	0.87	0.61	0.55	0.56
SVM	0.47	0.53	0.42	0.45
KNN	0.69	0.68	0.47	0.51

Berdasarkan hasil perbandingan tiga model klasifikasi yaitu Random Forest, SVM, dan KNN, terlihat bahwa terdapat perbedaan performa yang cukup jelas pada setiap metrik evaluasi.

Model Random Forest menunjukkan performa terbaik dengan akurasi sebesar 0,87, serta nilai precision 0,65, recall 0,60, dan F1-score 0,62. Hal ini menandakan bahwa model ini mampu mengklasifikasikan data dengan cukup baik dan seimbang dibandingkan model lainnya. Model KNN berada di posisi kedua dengan akurasi 0,69. Meskipun precision cukup tinggi yaitu 0,68, nilai recall yang lebih rendah (0,47) menunjukkan masih adanya data yang belum terdeteksi dengan baik, sehingga F1-score hanya mencapai 0,51. Sementara itu, model SVM memiliki performa terendah dengan akurasi 0,47, serta nilai precision 0,53, recall 0,42, dan F1-score 0,45, yang menunjukkan bahwa model ini kurang optimal dalam mengenali pola data.

Secara keseluruhan, Random Forest menjadi model paling unggul untuk dataset ini karena memiliki keseimbangan performa terbaik di antara ketiga model.

4.7 Uji ANOVA

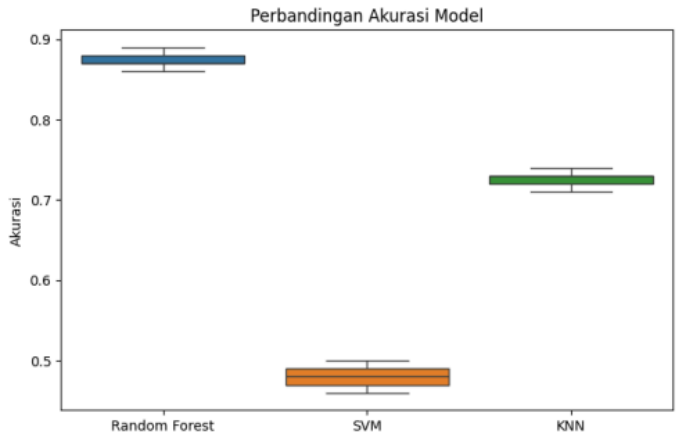
Uji ANOVA dilakukan untuk mengetahui apakah terdapat perbedaan yang signifikan antara performa model Random Forest, SVM, dan KNN.

Tabel 4. 14 Hasil Uji ANOVA Perbandingan Model

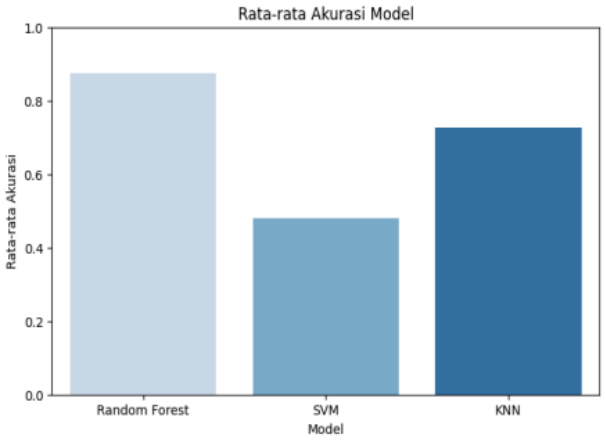
Metode	F-value	P-value	Keterangan
Random Forest, KNN, SVM	1164.98	< 0.001	Terdapat Perbedaan Signifikan

Berdasarkan hasil uji ANOVA pada model Random Forest, K-Nearest Neighbor (KNN), dan Support Vector Machine (SVM), diperoleh nilai F-value sebesar 1164,98 dan P-value sebesar < 0.001. Nilai P-value tersebut lebih kecil dari tingkat signifikansi 0,05, sehingga keputusan yang diambil adalah menolak H0. Hasil tersebut menunjukkan bahwa terdapat perbedaan performa yang signifikan antara model Random Forest, KNN, dan SVM dalam melakukan klasifikasi data. Nilai F-value yang sangat besar juga menunjukkan bahwa perbedaan performa antar model cukup tinggi dan tidak terjadi secara kebetulan.

Berdasarkan hasil pengujian sebelumnya, model Random Forest memiliki performa terbaik dibandingkan KNN dan SVM. Selain itu, Random Forest juga lebih stabil dalam mengklasifikasikan seluruh kelas data karena memiliki nilai evaluasi yang lebih seimbang. Dengan demikian, hasil uji ANOVA memperkuat bahwa terdapat perbedaan nyata pada performa ketiga model yang digunakan dalam penelitian ini.



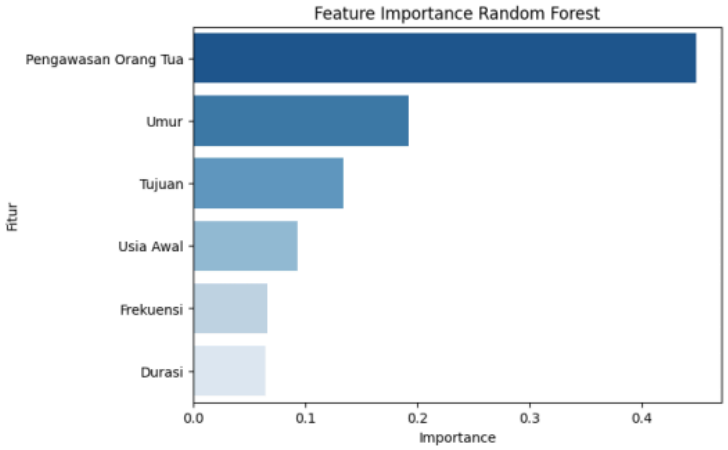
Gambar 4.11 Boxplot Hasil Uji ANOVA



Gambar 4.12 Rata-Rata Akurasi Ketiga Model

4.8 Analisis Feature Importance

Salah satu keunggulan Random Forest adalah kemampuannya dalam menentukan tingkat kepentingan masing-masing variabel (feature importance). Berdasarkan hasil analisis, diperoleh bahwa variabel yang paling berpengaruh terhadap risiko ketergantungan smartphone adalah:



Gambar 4.13 Diagram Feature Importance

Berdasarkan hasil analisis feature importance menggunakan algoritma Random Forest, diperoleh bahwa setiap variabel memiliki tingkat kontribusi yang berbeda dalam proses klasifikasi tingkat risiko ketergantungan smartphone pada

anak usia dini. Variabel yang memiliki pengaruh paling besar adalah Pengawasan Orang Tua dengan nilai importance sekitar 0,45. Hasil ini menunjukkan bahwa tingkat pengawasan yang diberikan orang tua merupakan faktor utama yang memengaruhi prediksi risiko ketergantungan smartphone pada anak.

Variabel Umur menempati urutan kedua dengan nilai importance sekitar 0,19. Hal ini mengindikasikan bahwa perbedaan usia anak turut berpengaruh terhadap tingkat risiko ketergantungan smartphone. Selanjutnya, variabel Tujuan Penggunaan Smartphone memiliki nilai importance sekitar 0,14, yang menunjukkan bahwa aktivitas yang dilakukan anak saat menggunakan smartphone juga berkontribusi terhadap hasil klasifikasi. Variabel Usia Awal Mengenal Smartphone memiliki nilai importance sekitar 0,09, sedangkan Frekuensi Penggunaan Smartphone dan Durasi Penggunaan Smartphone memiliki nilai importance yang lebih rendah, masing-masing sekitar 0,07 dan 0,06. Meskipun kedua variabel tersebut sering dianggap sebagai faktor utama dalam penggunaan smartphone, hasil penelitian menunjukkan bahwa pengaruhnya masih berada di bawah faktor pengawasan orang tua dan karakteristik anak.

Secara keseluruhan, hasil feature importance menunjukkan bahwa faktor yang berkaitan dengan peran dan pengawasan orang tua memiliki kontribusi paling dominan dalam menentukan tingkat risiko ketergantungan smartphone pada anak usia dini. Dengan demikian, peningkatan pengawasan dan pendampingan orang tua dapat menjadi salah satu upaya penting dalam mengurangi risiko ketergantungan smartphone pada anak.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, algoritma Random Forest berhasil digunakan untuk membangun model prediksi risiko ketergantungan smartphone pada anak usia dini berdasarkan enam variabel independen, yaitu umur, durasi penggunaan smartphone, frekuensi penggunaan smartphone, usia pertama kali mengenal smartphone, tujuan penggunaan smartphone, dan pola pengawasan orang tua. Model yang dibangun mampu mengklasifikasikan tingkat risiko ketergantungan smartphone berdasarkan data responden, sehingga dapat dimanfaatkan sebagai pendukung dalam mendeteksi risiko ketergantungan smartphone pada anak usia dini.

Hasil pengujian menunjukkan bahwa algoritma Random Forest memiliki performa yang lebih baik dibandingkan algoritma Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN). Random Forest memperoleh nilai accuracy sebesar 87%, precision sebesar 61%, recall sebesar 55%, dan F1-score sebesar 56%. Selain itu, hasil uji ANOVA menunjukkan adanya perbedaan performa yang signifikan antar algoritma, sehingga dapat disimpulkan bahwa Random Forest merupakan algoritma yang paling optimal untuk diterapkan dalam penelitian ini.

Berdasarkan hasil analisis *feature importance* pada algoritma Random Forest, variabel pola pengawasan orang tua merupakan faktor yang paling berpengaruh dalam memprediksi risiko ketergantungan smartphone pada anak usia dini. Temuan ini menunjukkan bahwa pengawasan orang tua memiliki peran

yang sangat penting dalam mengendalikan penggunaan smartphone pada anak. Dengan demikian, algoritma Random Forest tidak hanya mampu memberikan hasil prediksi dengan performa yang baik, tetapi juga dapat mengidentifikasi faktor-faktor yang paling berkontribusi terhadap risiko ketergantungan smartphone sehingga dapat mendukung upaya deteksi dan pencegahan sejak dini.

5.2 Saran

Berdasarkan keterbatasan penelitian, penelitian selanjutnya disarankan menggunakan jumlah responden yang lebih banyak dengan cakupan wilayah yang lebih luas agar hasil penelitian memiliki tingkat generalisasi yang lebih baik. Selain itu, penelitian dapat menambahkan variabel lain yang berpotensi memengaruhi risiko ketergantungan smartphone pada anak usia dini, seperti tingkat pendidikan orang tua, kondisi sosial ekonomi keluarga, dan lingkungan anak, sehingga model prediksi yang dihasilkan menjadi lebih akurat.

Penelitian selanjutnya juga disarankan membandingkan Random Forest dengan algoritma machine learning lainnya serta menggunakan metode evaluasi statistik yang lebih beragam untuk memperoleh perbandingan performa model yang lebih komprehensif. Selain itu, model prediksi yang dihasilkan dapat dikembangkan menjadi sistem pendukung keputusan atau aplikasi deteksi dini yang dapat membantu orang tua dan pihak terkait dalam mengidentifikasi risiko ketergantungan smartphone pada anak usia dini.

DAFTAR PUSTAKA

[1] A. L. B. Masalah, "Dikdik M. Arief Mansur, Cyber Law: Aspek Hukum Teknologi Informasi, Ed. Oleh Pt Rapfika Aditama, Pt Rapfika Aditama (Bandung, 2005).121 1 1," Pp. 1–16.

[2] A. C. Fatimah, "Peran Orang Tua Dalam Mengoptimalisasi Ketergantungan Anak Pada Smartphone," *Fitrah J. Stud. Pendidik.*, Vol. 12, No. 2, Pp. 53–74, 2022, Doi: 10.47625/Fitrah.V12i2.337.

[3] M. A. Eva Manfaatin, "Jspa : Jurnal Terdiri Aplikasinya," *J. Pendidik. Islam Anak Usia Dini*, Vol. 01, No. 01, P. 18, 2024.

[4] D. Arman *Et Al.*, "Jurnal Software Engineering And Information System (Seis) Pemodelan Machine Learning Dengan Algoritma Random Forest," Vol. 5, No. 2, 2025.

[5] K. Ilmiah, "Jurnal Edu Research Indonesian Institute For Corporate Learning And Studies (Iicls) Page 90," Vol. 5, No. September, Pp. 90–97, 2024.

[6] P. Smartphone, T. Anak, And U. Dini, "Pengaruh Smartphone Terhadap Anak Usia Dini," Vol. 3, No. 2, 2023.

[7] P. Studi, S. Informasi, And U. I. Indragiri, "Rancang Bangun Sistem Prediksi Pola Sidik," Vol. 11, No. 3, 2023.

[8] N. F. Sahamony And H. Rianto, "Analysis Of Performance Comparison Of Machine Learning Models For Predicting Stunting Risk In Children ' S Growth Analisis Perbandingan Kinerja Model Machine Learning Untuk Memprediksi Risiko Stunting Pada Pertumbuhan Anak," Vol. 4, No. April, Pp. 413–422, 2024.

[9] M. Fauzan *Et Al.*, "Fenomena Kecanduan Gadget Pada Anak Usia Pra Sekolah

Sosial Perilaku Kecanduan Anak Pra-Sekolah Faktor Lingkungan . Menurut Data Kementrian Penyelenggara Jasa Internet Indonesia (Apjii) Memiliki Total Jumlah Penduduk Sebanyak 282 Riset Yang Dilakukan Pada Tahun 2014 Oleh The Asian Parent Insights Menyatakan Terdapat," Vol. 5, No. 3, 2024.

[10] S. Regiana, L. Suzanti, And D. Widjayatri, "Penggunaan Gadget Pada Anak Usia 4-5 Tahun Dalam Keluarga Dengan Pola Asuh Permisif," Vol. 5, No. 1, Pp. 446–458, 2024, Doi: 10.37985/Murhum.V5i1.575.

[11] M. Maulidah And N. Hidayati, "Prediksi Kesehatan Tidur Dan Gaya Hidup," Vol. 4, No. 1, Pp. 81–86, 2024.

[12] B. I. Nugroho, N. A. Santoso, A. A. Murtopo, P. Korespondensi, And I. Pendahuluan, "Prediksi Kemampuan Akademik Mahasiswa Dengan Metode Support Vector Machine," Vol. 7, Pp. 177–188, 2023.

[13] M. Kwon *Et Al.*, "Development And Validation Of A Smartphone Addiction Scale (Sas)," Vol. 8, No. 2, 2013, Doi: 10.1371/Journal.Pone.0056936.

[14] E. Katz, J. A. Y. G. Blumler, And M. Gurevitch, "Uses And Gratifications Research," *Public Opin. Q.*, Vol. 37, No. 4, Pp. 509–523, Jan. 1973, Doi: 10.1086/268109.

[15] J. R. Marlatt, G. Alan; Gordon, *Relapse Prevention: Maintenance Strategies In The Treatment Of Addictive Behaviors*. New York: Guilford Press, 1985.

[16] P. M. Valkenburg, M. Krcmar, A. L. Peeters, And N. M. Marseille, "Developing A Scale To Assess Three Styles Of Television Mediation: 'Instructive Mediation,' 'Restrictive Mediation,' And 'Social Coviewing,'" *J. Broadcast. Electron. Media*, Vol. 43, No. 1, Pp. 52–66, Jan. 1999, Doi:

10.1080/08838159909364474.

[17] T. M. Mitchell, *Machine Learning*. New York: Mcgraw-Hill, 1997.

[18] E. Alpaydin, *Introduction To Machine Learning*, 4th Editio. Cambridge, Massachusetts: Mit Press, 2020.

[19] L. E. O. Breiman, "Random Forests," Pp. 5–32, 2001.

[20] H. Han, Jiawei; Pei, Jian; Tong, *Data Mining: Concepts And Techniques*, 4th Ed. Cambridge, Ma, Usa: Morgan Kaufmann (An Imprint Of Elsevier), 2022.

[21] A. R. Supriyatna *Et Al.*, "Prediksi Penyakit Diabetes Menggunakan Algoritma Random Forest," Vol. 17, No. X, Pp. 163–172, 1978.

[22] N. H. Alfajr And S. Defiyanti, "Metode Random Forest Dan Penerapan Principal Component Analysis (Pca)," Vol. 12, No. 3, 2024.

SKRIPSI Kamia Pertiwi Sistem Informasi.pdf

ORIGINALITY REPORT

40%
SIMILARITY INDEX

PRIMARY SOURCES

1	Kamia Pertiwi, Usman Usman, Ilyas Ilyas. "Prediksi Risiko Ketergantungan Smartphone Pada Anak Usia Dini Menggunakan Algoritma Random Forest", Indonesian Journal of Multidisciplinary on Social and Technology, 2026 Crossref	2131 words — 16%
2	repository.unisi.ac.id Internet	289 words — 2%
3	repository.uin-suska.ac.id Internet	113 words — 1%
4	repository.machung.ac.id Internet	108 words — 1%
5	proceeding.unpkediri.ac.id Internet	104 words — 1%
6	eprints.amikom.ac.id Internet	65 words — < 1%
7	repository.upi.edu Internet	64 words — < 1%
8	journal.jisti.unipol.ac.id Internet	57 words — < 1%
9	repo.darmajaya.ac.id Internet	55 words — < 1%
10	kc.umn.ac.id Internet	54 words — < 1%
11	jurnal.polibatam.ac.id Internet	44 words — < 1%

12	Luthfia Nurma Hapsari, Ilham Fannani, Yenny Rahmawati, Ahmad Muhariya. "Explainable Machine Learning untuk Prediksi Risiko Penyakit Jantung Menggunakan Random Forest dan Analisis SHAP", remik, 2026 Crossref	42 words — < 1%
13	Akbar Supia Dirja, Lana Aulia Salsabila, M. Akhdan Putra Hermawan, Cika Jelita, Hanif Zaidan Sinaga. "Analisis Sentimen MBG Di Media Sosial X Menggunakan Support Vector Machine Dan Naive Bayes", Indonesian Journal of Multidisciplinary on Social and Technology, 2026 Crossref	40 words — < 1%
14	dirdosen.budiluhur.ac.id Internet	40 words — < 1%
15	jurnal.fikom.umi.ac.id Internet	40 words — < 1%
16	repository.ub.ac.id Internet	40 words — < 1%
17	Muhammad Suhada, Adil Setiawan, Roslina Roslina. "Perbandingan Kinerja Metode Random Forest dan Logistic Regression Dalam Konteks Prediksi Penerima Beasiswa Bagi Mahasiswa Baru", Jurnal Minfo Polgan, 2026 Crossref	39 words — < 1%
18	repository.ar-raniry.ac.id Internet	35 words — < 1%
19	123dok.com Internet	34 words — < 1%
20	jurnal.darmajaya.ac.id Internet	34 words — < 1%
21	repository.pnj.ac.id Internet	33 words — < 1%
22	text-id.123dok.com Internet	33 words — < 1%
23	kanal.psikologi.ugm.ac.id Internet	32 words — < 1%

24	edu.pubmedia.id Internet	30 words — < 1%
25	repository.unhas.ac.id Internet	29 words — < 1%
26	www.scribd.com Internet	28 words — < 1%
27	Rena Nurika Salsabila, Laela Lutfiana Rachmah, Dessy Farantika. "Bentuk Dan Makna Pendampingan Orang Tua Terhadap Penggunaan Smartphone Anak Usia Dini Berdasarkan Parental Mediation Theory (Studi Kasus Di Kelurahan Kauman)", Science and Education Journal (SICEDU), 2026 Crossref	27 words — < 1%
28	Ronatio, Januar. "Pengaruh Metode Sorogan dalam Pembelajaran Ilmu Tajwid terhadap Kemampuan Membaca Al-Qur'an pada Siswa di SMA N I X Koto Kabupaten Tanah Datar", YASIN, 2026 Crossref	27 words — < 1%
29	repository.unilak.ac.id Internet	27 words — < 1%
30	repository.unmuhjember.ac.id Internet	26 words — < 1%
31	es.scribd.com Internet	25 words — < 1%
32	id.123dok.com Internet	25 words — < 1%
33	Bangkit Indarmawan Nugroho, Nugroho Adhi Santoso, Aang Alim Murtopo. "Prediksi Kemampuan Akademik Mahasiswa dengan Metode Support Vector Machine", remik, 2023 Crossref	24 words — < 1%
34	etheses.uin-malang.ac.id Internet	24 words — < 1%
35	journal.stmikjayakarta.ac.id Internet	24 words — < 1%
36	jurnal.unimor.ac.id Internet	

		24 words — < 1%
37	lib.unnes.ac.id Internet	24 words — < 1%
38	Widi Sriastuti. "Pengaruh Pola Asuh Orang Tua terhadap Screen Time pada Anak Usia Dini (di RA Al-Abror Kecamatan Kersamanah Kabupaten Garut)", AS-SABIQUN, 2025 Crossref	23 words — < 1%
39	Yuce Yuliani. "Analisis Komparatif Algoritma Decision Tree dan Random Forest untuk Prediksi Kelulusan Mahasiswa", RIGGS: Journal of Artificial Intelligence and Digital Business, 2026 Crossref	23 words — < 1%
40	jurnal.unpad.ac.id Internet	23 words — < 1%
41	repository.radenintan.ac.id Internet	23 words — < 1%
42	repository.unissula.ac.id Internet	23 words — < 1%
43	www.coursehero.com Internet	23 words — < 1%
44	journal.unimma.ac.id Internet	22 words — < 1%
45	Siti Hajar, Nurdin Nurdin. "PERBANDINGAN KLASIFIKASI ALGORITMA DECISION TREE DAN RANDOM FOREST PADA DATASET INDEX PRESTASI SEMESTER MAHASISWA (IPS)", JUTECH : Journal Education and Technology, 2026 Crossref	21 words — < 1%
46	dspace.uii.ac.id Internet	21 words — < 1%
47	digilibadmin.unismuh.ac.id Internet	20 words — < 1%
48	media.neliti.com Internet	19 words — < 1%

49	ojs.udb.ac.id Internet	19 words — < 1%
50	repositori.unsil.ac.id Internet	19 words — < 1%
51	Antonius Andri Geong. "Analisis Distribusi Umur Pengelola Usaha Pertanian Perorangan (UTP) di Provinsi Nusa Tenggara Timur", Jurnal Statistika Terapan (ISSN 2807-6214), 2025 Crossref	18 words — < 1%
52	Haikam Mawazir, Ahmad Ipmawan Kharisma, Oriza Zativalen. "Hubungan Intensitas Penggunaan Smartphone Dengan Keterampilan Sosial-Emosional Siswa Sekolah Dasar", Science and Education Journal (SICEDU), 2026 Crossref	18 words — < 1%
53	Selvi Nuriana, Fitri Yunita. "Prediksi Stunting Menggunakan Random Forest dan SMOTE Berdasarkan Data Posyandu", RIGGS: Journal of Artificial Intelligence and Digital Business, 2026 Crossref	18 words — < 1%
54	eprints.unpak.ac.id Internet	18 words — < 1%
55	journal.stieamkop.ac.id Internet	18 words — < 1%
56	murhum.ppjpaud.org Internet	18 words — < 1%
57	repository.uinjambi.ac.id Internet	18 words — < 1%
58	repository.ulb.ac.id Internet	18 words — < 1%
59	Martin Martin, Wahyuda Dwi Handoko, Riki Maulana. "Analisis hubungan intensitas penggunaan gawai (screen time) dengan perkembangan sosial-emosional anak usia dini pada era digital", Jurnal Pendidikan Informatika dan Sains, 2024 Crossref	17 words — < 1%

60	Muhammad Wisnu Nugroho. "Analisis Performa Algoritma Random Forest dalam Mengatasi Overfitting pada Model Prediksi", Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi), 2025 Crossref	17 words — < 1%
61	eprints.upj.ac.id Internet	17 words — < 1%
62	journal.eng.unila.ac.id Internet	17 words — < 1%
63	juris.id Internet	17 words — < 1%
64	jurnal.umb.ac.id Internet	17 words — < 1%
65	repository.umsu.ac.id Internet	17 words — < 1%
66	Riska Aini Putri, Muhammad Abdul Hadi, Fuad Maulana, Dzaky Akram, Nathasa Septianingrum, Gunawan Gunawan. "LetsGrow Health: Prediksi Risiko Stunting pada Anak Menggunakan Algoritma Random Forest Berbasis Data WHO", RIGGS: Journal of Artificial Intelligence and Digital Business, 2026 Crossref	16 words — < 1%
67	cibinstitute.id Internet	16 words — < 1%
68	repo.ukitoraja.ac.id Internet	16 words — < 1%
69	repository.uinsaizu.ac.id Internet	16 words — < 1%
70	Marwa Sulehu, Wisda Wisda, First Wanita, Markani Markani. "Optimasi Prediksi Kelulusan Mahasiswa Menggunakan Random Forest untuk Meningkatkan Tingkat Retensi", Jurnal Minfo Polgan, 2025 Crossref	14 words — < 1%
71	Nur Lailiyah, Yolanda Rensia Gigik, Avrillya Ainisya, Salma Kurnia Anisa, Mutiara Etikasari. "Implementation of the Role-Playing Model	14 words — < 1%

in Indonesian Language Learning and Its Impact on Communication Skills", Efektor, 2025
Crossref

72	digilib.uin-suka.ac.id Internet	14 words — < 1%
73	ejurnal.seminar-id.com Internet	14 words — < 1%
74	eprints.ubhara.ac.id Internet	14 words — < 1%
75	media.hadlacorp.com Internet	14 words — < 1%
76	pmc.ncbi.nlm.nih.gov Internet	14 words — < 1%
77	readmore.id Internet	14 words — < 1%
78	Laila Salsabilla Hanifa, Herlawati Herlawati, Rakhmat Purnomo. "Perbandingan Kinerja Random Forest dan Long Short-Term Memory dalam Prediksi Kedalaman Hiposentrum Gempa di Maluku dan Laut Banda", Journal of Students' Research in Computer Science, 2026 Crossref	13 words — < 1%
79	eprints.bsi.ac.id Internet	13 words — < 1%
80	journal.fkpt.org Internet	13 words — < 1%
81	lab.tik.pnj.ac.id Internet	13 words — < 1%
82	multidisipliner.org Internet	13 words — < 1%
83	ojs.dcckotabumi.ac.id Internet	13 words — < 1%
84	repository.ubb.ac.id Internet	13 words — < 1%

85	Hestia Alika K, Handriani J.H, Lisa Savatri Nataleni, Salini Salini, Selvia Veronika, Wirastiani Binti Yusup. "Pemanfaatan Media Pembelajaran Puzzle dalam Meningkatkan Kemampuan Kognitif Anak Usia Dini di TK Marina Permai", Jurnal Pendidikan Anak Usia Dini, 2025 Crossref	12 words — < 1%
86	ejournal.iain-manado.ac.id Internet	12 words — < 1%
87	jurnaldashd.wordpress.com Internet	12 words — < 1%
88	repository.syekhnurjati.ac.id Internet	12 words — < 1%
89	repository.unas.ac.id Internet	12 words — < 1%
90	repository.untad.ac.id Internet	12 words — < 1%
91	Andreyestha Andreyestha, Agus Subekti. "ANALISA SENTIMENT PADA ULASAN FILM DENGAN OPTIMASI ENSEMBLE LEARNING", Jurnal Informatika, 2020 Crossref	11 words — < 1%
92	Azzumardi Mubarok Romadhoni, Widyasari Puspa Permata Witra, Indra Gita Anugrah. "PENERAPAN ALGORITMA RANDOM FOREST SEBAGAI EARLY WARNING SYSTEM PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU", INTECOMS: Journal of Information Technology and Computer Science, 2026 Crossref	11 words — < 1%
93	Putri Cholidhazia, Pedro Masdy, Arif Mudi Priyatno. "Early Prediction of Perinatal Depression Risk Among Pregnant Women Using Random Forest and XGBoost Algorithms", Indonesian Journal of Multidisciplinary on Social and Technology, 2026 Crossref	11 words — < 1%
94	ejurnal.provisi.ac.id Internet	11 words — < 1%
95	journal.ilmudata.co.id Internet	11 words — < 1%

96	journal.universitaspahlawan.ac.id Internet	11 words — < 1%
97	jurnalpariwisata.iptrisakti.ac.id Internet	11 words — < 1%
98	repositori.usu.ac.id Internet	11 words — < 1%
99	repository.unpas.ac.id Internet	11 words — < 1%
100	Md Wira Putra Dananjaya, Gede Humaswara Prathama, I Gusti Ngurah Darma Paramartha, Putu Gita Pujayanti. "Klasifikasi Kualitas Tanah Berdasarkan Kandungan pH, Kelembapan, dan Suhu Menggunakan Algoritma K-Nearest Neighbors", Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi), 2025 Crossref	10 words — < 1%
101	Raras Tyasnurita, Adhi Yoga Muris Pamungkas. "Deteksi Diabetik Retinopati menggunakan Regresi Logistik", ILKOM Jurnal Ilmiah, 2020 Crossref	10 words — < 1%
102	core.ac.uk Internet	10 words — < 1%
103	e-journal.hamzanwadi.ac.id Internet	10 words — < 1%
104	eprints.uns.ac.id Internet	10 words — < 1%
105	eprints.walisongo.ac.id Internet	10 words — < 1%
106	journal.poltekad.ac.id Internet	10 words — < 1%
107	jurnal.alimspublishing.co.id Internet	10 words — < 1%
108	mfile.narotama.ac.id Internet	10 words — < 1%

109	ojs.stieamkop.ac.id Internet	10 words — < 1%
110	peerj.com Internet	10 words — < 1%
111	repository.teknokrat.ac.id Internet	10 words — < 1%
112	repository.upnvj.ac.id Internet	10 words — < 1%
113	www.journal.irpi.or.id Internet	10 words — < 1%
114	www.researchgate.net Internet	10 words — < 1%
115	Nadiya Anisa Pratidina, Ahmad Suriansyah, Wahdah Refia Rafianti. "Penggunaan Smartphone dalam Pembelajaran di Sekolah Dasar", MARAS: Jurnal Penelitian Multidisiplin, 2024 Crossref	9 words — < 1%
116	Wan Dinulaqli, Nurdin Nurdin. "PERBANDINGAN PERFORMA NAÏVE BAYES DAN MODIFIKASI NAIVE BAYES DENGAN LAPLACE CORRECTION PADA KLASIFIKASI POTENSI BANJIR MENGGUNAKAN DATA CURAH HUJAN", JUTECH : Journal Education and Technology, 2026 Crossref	9 words — < 1%
117	eprints.iain-surakarta.ac.id Internet	9 words — < 1%
118	eprints.uny.ac.id Internet	9 words — < 1%
119	geograf.id Internet	9 words — < 1%
120	journals.upi-yai.ac.id Internet	9 words — < 1%
121	ojs.staira.ac.id Internet	9 words — < 1%

122	paud-anakbermainbelajar.blogspot.com Internet	9 words — < 1%
123	repository.unjaya.ac.id Internet	9 words — < 1%
124	repository.unmuha.ac.id:8080 Internet	9 words — < 1%
125	Achmad Hakim Qoirul Haq, Harminto Mulyo, Adi Sucipto. "Optimasi Kinerja Algoritma K-Nearest Neighbor melalui Metode Random Forest untuk Klasifikasi Penyakit Ginjal", Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi), 2026 Crossref	8 words — < 1%
126	Athallah Yasyfi Imran, M. Rudi Sanjaya, Bayu Wijaya Putra, Gabriel Ekoputra Hartono Cahyadi. "Optimization of Sentiment Analysis on Tokopedia User Reviews Using Gridsearchcv and Smote with Machine Learning Algorithms", INOVTEK Polbeng - Seri Informatika, 2025 Crossref	8 words — < 1%
127	Lely Kurniawati, Dadang Priyanto, Neny Sulistia Ningsih, Moch Syahrir, Ria Rismayati. "Comparison of Decision Tree-Based Methods in Lung Disease Detection", Jurnal Bumigora Information Technology (BITe), 2025 Crossref	8 words — < 1%
128	Maulana Fansyuri, Devi Yunita. "Analisis Kinerja Algoritma K-Nearest Neighbor dalam Prediksi Kelulusan Berdasarkan Data Akademik Menggunakan Python", Indonesian Journal of Multidisciplinary on Social and Technology, 2026 Crossref	8 words — < 1%
129	Rahma Cahya Pratiwi, Aprilia Anggun Setyo Budi. "Technoference dalam Pola Asuh Digital: Studi Fenomenologi tentang Interaksi Emosional Orang Tua dan Anak Usia Sekolah Dasar di Kota Yogyakarta", JURNAL PENDIDIKAN IPS, 2026 Crossref	8 words — < 1%
130	Rinda Fitriana Azmi, Galih Hendro Martono, Neny Suliastianingsih. "Komparasi Algoritma Random Forest dan XGBoost dalam Klasifikasi Risiko Perizinan Usaha pada UMKM", RIGGS: Journal of Artificial Intelligence and Digital Business, 2026 Crossref	8 words — < 1%

131	Serli Marlina, Rismareni Pransiska, Zahratul Qalbi. "Analisis Kurikulum Pendidikan Islam di Taman Kanak-Kanak Aisyiyah Bustanul Athfal Padang", Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini, 2021 Crossref	8 words — < 1%
132	bajangjournal.com Internet	8 words — < 1%
133	digilib.uinkhas.ac.id Internet	8 words — < 1%
134	docplayer.info Internet	8 words — < 1%
135	ejournal.nusamandiri.ac.id Internet	8 words — < 1%
136	ejurnal.ung.ac.id Internet	8 words — < 1%
137	etd.repository.ugm.ac.id Internet	8 words — < 1%
138	etheses.iainponorogo.ac.id Internet	8 words — < 1%
139	hostjournals.com Internet	8 words — < 1%
140	journal.sinov.id Internet	8 words — < 1%
141	jurnal.delitekno.co.id Internet	8 words — < 1%
142	kishi-kun.blogspot.com Internet	8 words — < 1%
143	pinkylovers-braille.blogspot.com Internet	8 words — < 1%
144	pustakabakul.blogspot.com Internet	8 words — < 1%

145	repository.umnaw.ac.id Internet	8 words — < 1%
146	repository.unisma.ac.id Internet	8 words — < 1%
147	repository.unsoed.ac.id Internet	8 words — < 1%
148	sistemasi.ftik.unisi.ac.id Internet	8 words — < 1%
149	www.journal.unpas.ac.id Internet	8 words — < 1%
150	Inggarwangi Dinda Praditasari, Zainal Abidin, Nurkinan Nurkinan. "PEMENUHAN KEBUTUHAN INFORMASI PENGGEMAR BTS (ARMY) MELALUI AKUN INSTAGRAM @COPPAMAGZ", Network Media, 2026 Crossref	7 words — < 1%
151	Melania Velisita Lau, Aryani Dwi Handayani, Miyosagusti Yudo Widodo, Rani Irma Handayani, Risca Lusiana Pratiwi. "Implementasi Metode Random Forest untuk Klasifikasi Sampah Organik dan Anorganik Berbasis Citra Digital", RIGGS: Journal of Artificial Intelligence and Digital Business, 2025 Crossref	7 words — < 1%
152	Muhammad Fitra Ramadhan, Prastiya Indra Gunawa, Muhammad Arifin Parenrengi. "Faktor-Faktor Risiko Pasien Epilepsi Intraktabel pada Anak", Malahayati Nursing Journal, 2022 Crossref	7 words — < 1%
153	Rahmawati, Nur Laili. "Efektifitas Kombinasi Model Pembelajaran Kooperatif Tipe Think Pair Share (TPS) Dan Pendekatan Open Ended Untuk Meningkatkan Ketrampilan Siswa Dalam Memecahkan Masalah Matematika.", Universitas Islam Negeri Saifuddin Zuhri (Indonesia) ProQuest	7 words — < 1%
154	Reviza Mayhart, Suryanto Suryanto. "Dampak Technoference dan Parental Phubbing terhadap Perkembangan Sosial-Emosional Anak Usia Dini: Systematic Literature Review", Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini, 2026 Crossref	7 words — < 1%

155	Abdullah Abdullah, Zikrullah Zikrullah, Usman Usman. "RANCANG BANGUN SISTEM PREDIKSI POLA SIDIK JARI MENGGUNAKAN METODE BACKPROPAGATION", Jurnal Informatika dan Teknik Elektro Terapan, 2023 Crossref	6 words — < 1%
156	Fiki Febrian Dwi Prasetya. "PSIKOEDUKASI PENCEGAHAN KECANDUAN GADGET USIA ANAK-ANAK PADA TPQ HALIM RAMADHAN TUBO", Al-Khidmah : Jurnal Pengabdian dan Pendampingan Masyarakat, 2026 Crossref	6 words — < 1%
157	Gulnoria, Revy Amyneva. "Pengaruh Akuntabilitas Islam dan Transparansi Terhadap Tingkat Kepercayaan Muzakki Dengan Good Zakat Governance Sebagai Variabel Intervening pada Badan Amil Zakat Nasional Provinsi Jawa Tengah", Universitas Islam Sultan Agung (Indonesia), 2024 ProQuest	6 words — < 1%
158	Sarman, Sarman. "Pembentukan Akhlak Karimah Pada Siswa Smp Ma'had Darussa'adah Gununglurah Cilongok Di Kabupaten Banyumas", Institut Agama Islam Negeri Purwokerto (Indonesia), 2022 ProQuest	6 words — < 1%
159	Siswanto, Siswanto. "Peran Orang Tua Dalam Pendidikan Ibadah Anak Di Desa Rejosari Kecamatan Karangawen Kabupaten Demak", Universitas Islam Sultan Agung (Indonesia), 2022 ProQuest	6 words — < 1%
160	Thazra Nur Arifah, Ikhwan Hamdani, Syarifah Gustiawati. "Pengaruh Kualitas Produk dan Pemasaran Islami melalui WhatsApp terhadap Kepuasan Konsumen UMKM Kerupuk Ibu Tarsih", AHKAM, 2025 Crossref	6 words — < 1%
161	repository.unj.ac.id Internet	6 words — < 1%
162	zombiedoc.com Internet	6 words — < 1%

EXCLUDE QUOTES	OFF	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	OFF