Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri

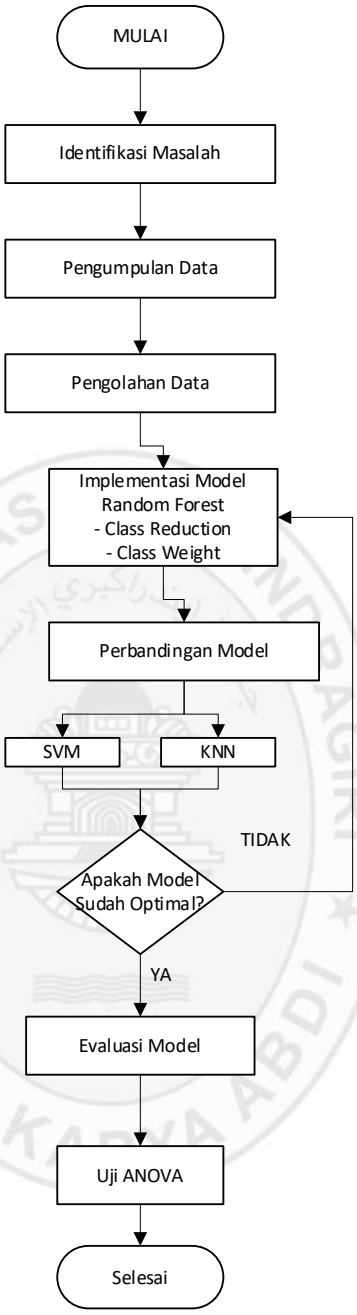
BAB III

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

Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri

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.

Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri

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.

Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri

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.

Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri

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:

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

Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri

Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri

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.*

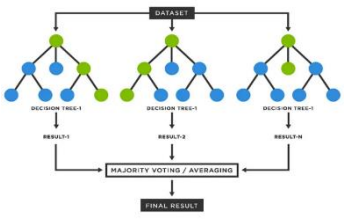
Jika p-value ≥ 0.05 → *Tidak terdapat perbedaan signifikan.*

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

Hak Cipta Dilindungi Undang-Undang
1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.
Universitas Islam Indragiri