

## IDENTIFIKASI PAPARAN RADIASI X-RAY UNTUK KESELAMATAN RADIASI MENGGUNAKAN *RANDOM FOREST CLASSIFICATION*

Diah Rahayu Ningtias<sup>1\*</sup>, Muhammad Rofi'i<sup>1</sup>, Bayu Wahyudi<sup>1</sup>, Josepa ND Simanjuntak<sup>2</sup>, Rodhotul Muttaqin<sup>3</sup>

<sup>1</sup>Program Studi Teknik Elektromedik, Sekolah Tinggi Ilmu Kesehatan Semarang, Semarang, Indonesia

<sup>2</sup>Departemen Radiologi, Rumah Sakit Umum Pusat Adam Malik, Medan, Indonesia

<sup>3</sup>Laboratorium Fisika, Universitas Negeri Semarang, Semarang, Indonesia

\*Email: [diahrahayuningtias@stikessemarang.ac.id](mailto:diahrahayuningtias@stikessemarang.ac.id)

### ABSTRAK

Tingkat keamanan paparan X-Ray sangat penting karena mempunyai dampak jangka pendek dan jangka panjang yang sangat mempengaruhi kesehatan pekerja radiasi dan lingkungan sekitarnya. Pengukuran tingkat keamanan radiasi X-Ray pada umumnya dilakukan dengan cara konvensional, yaitu mengidentifikasi secara manual data nilai paparan radiasi bagi pekerja dan lingkungan sekitar dari *survey* meter, kemudian dijumlahkan secara berkala. Namun hal ini berpotensi menimbulkan kesalahan dalam penambahan sehingga metode menghasilkan data yang kurang akurat. Penelitian ini bertujuan untuk mengetahui dosis paparan radiasi X-Ray menggunakan *Random Forest Classification*. Data radiasi yang diolah merupakan nilai dosis pengukuran paparan X-Ray dengan menggunakan *survey* meter digital (dalam  $\mu\text{Sv/h}$ ) sebanyak 160 *data set* dan terdiri dari 87 dosis aman dan 73 dosis tidak aman. Data diklasifikasikan menurut aturan nilai batas dosis *International Atomic Energy Agency* (IAEA). Kinerja *Random Forest Classification* dievaluasi dengan *Naïve Bayes* dan *K-Nearest Neighbor* (KNN). Hasil penelitian menunjukkan nilai akurasi *Random Forest Classification* sebesar 90%, nilai akurasi *Naïve Bayes* sebesar 85%, dan nilai akurasi KNN sebesar 86%. Oleh karena itu, nilai kinerja dari *Random Forest Classification* sebesar 97% diambil sebagai hasil terbaik. Sebagai rangkuman penelitian ini, *Random Forest Classification* berkinerja lebih baik dibandingkan *Naïve Bayes* dan *K-Nearest Neighbor* (KNN) lainnya dalam mengidentifikasi tingkat keamanan paparan radiasi X-Ray yang terbukti secara optimal berdasarkan parameter yang diterapkan.

**Kata kunci:** dosis radiasi; proteksi radiasi; detektor isian gas; identifikasi; akurasi

### ABSTRACT

*The safety level of X-ray exposure is very important because it has short-term and long-term effects that significantly affect the health of radiation workers and the surrounding environment. Measurements of the safety level of X-ray radiation are generally carried out using conventional methods, namely manually identifying the radiation exposure value data for workers and the surrounding environment from a survey meter, then adding up periodically. However, this could cause errors in the addition so that the method produces less accurate data. This study aims to X-ray radiation exposure dose using Random Forest Classification. The radiation data processed is the dose value of X-ray exposure measures using a digital survey meter (in  $\mu\text{Sv/h}$ ) unit of as many as 160 datasets and consists of 87 safe and 73 unsafe doses. Data are classified according to the International Atomic Energy Agency (IAEA) dosage limit value rule. The performance of Random Forest Classification is evaluated with Naïve Bayes dan K-Nearest Neighbor (KNN). The result shows that the Random Forest Classification accuracy value is 90%, the Naïve Bayes accuracy value is 85%, and the KNN accuracy value is 86%. Therefore, the performance value from the Random Forest Classification of 97% is the best result. As a summary of this study, Random Forest Classification performed better than other Naïve Bayes and K-Nearest Neighbor (KNN) for identifying the safety level of X-ray radiation exposure as proven with the optimum given the parameters applied.*

**Keywords:** radiation dose; radiation protection; gas fill detector; identification; accuracy

## 1. PENDAHULUAN

Salah satu kegunaan radiasi pengion dalam bidang medis adalah *X-Ray*. *X-Ray* diproduksi dalam tabung dioda vakum (katoda-anoda) yang disuplai ke sumber listrik tegangan tinggi (dalam kV). Ketika katoda diberi tegangan tinggi, maka memicu pembentukan elektron pada tingkat energi tertentu. Elektron tersebut melewati jalur bebas dalam tabung vakum, melepaskan elektron, dan kemudian mengenai anoda (target). Ketika transfer energi ini terjadi, karakteristik *X-Ray* terbentuk [1].

Dampak yang ditimbulkan oleh radiasi *X-Ray* antara lain kerusakan struktur molekul, seperti perubahan warna, perubahan kimia, dan kerusakan DNA pada sampel biologis. Kerusakan ini terjadi bila *X-Ray* yang dihasilkan mengalami serapan fotolistrik <30 keV. Tingkat energi ini menghasilkan foton atau fotoelektron, yang mengakibatkan proses yang dilakukan menjadi terlalu panas. Akibatnya, *X-Ray* berenergi tinggi umumnya lebih merusak dibandingkan *X-Ray* berenergi rendah [2]. Efek lain yang ditimbulkan oleh *X-Ray* berenergi tinggi adalah dosis yang diterima oleh pasien (benda) dan lingkungan sekitar [3]. Pengaruh dosis radiasi pada tubuh berbeda-beda menurut jenis dan tingkat energi yang diserap. Dosis setara diukur dari mikrosievert ( $\mu\text{Sv}$ ) hingga milisievert (mSv) dan digunakan untuk menentukan risiko terhadap kesehatan pasien. Nilai dosis setara di atas 200 mSv telah terbukti meningkatkan risiko kanker jika terpapar pada jaringan lunak. Namun, nilai tersebut masih dalam batas aman terhadap paparan tulang [4].

Nilai batas dosis untuk lingkungan yang tepat menurut *International Atomic Energy Agency* (IAEA) tidak lebih dari  $1 \mu\text{Sv/h}$  pada jarak 1 m dari sumber radiasi [5]. Ketebalan bahan pelindung mempengaruhi distribusi dosis di lingkungan sekitar sumber *X-Ray*. Standar yang dianjurkan untuk pekerja adalah  $10 \mu\text{Sv/h}$ , sedangkan untuk lingkungan sekitar  $0,5 \mu\text{Sv/h}$ . Nilai dosis ekuivalen ini diperoleh dari variasi faktor pemaparan tabung *X-Ray* [6]. Paparan radiasi dan dosis efektif mempunyai satuan yang sama dengan dosis serapan. Dosis paparan radiasi atau nilai batas dosis ekuivalennya bagi pengguna masyarakat berbeda dengan yang dianjurkan bagi pekerja radiasi. Pada penelitian ini, nilai batas standar dosis ekuivalen untuk pengguna masyarakat pada kulit adalah 50 mSv/tahun atau sama dengan  $5,707 \mu\text{Sv/h}$  [7]. Dosis pada sumber radiasi mempengaruhi kualitas gambar yang dihasilkan, yang dikenal dengan aktivitas pengendalian kualitas gambar. Beberapa parameter yang digunakan untuk menentukan kualitas gambar *X-Ray* meliputi tingkat terpasang, keseragaman, akurasi, resolusi dan resolusi spasial [8].

Penentuan besarnya paparan radiasi dari sumbernya dapat dilakukan secara kuantitatif dengan menggunakan sistem deteksi [9]. Identifikasi dosis radiasi pada *X-Ray* hendaknya membantu tenaga kesehatan dalam menggunakan sumber radiasi pengion khususnya *X-Ray*, serta lebih memperhatikan tingkat keamanan ruang rontgen agar tidak membahayakan lingkungan dan penggunanya. Penelitian ini juga dapat membantu menetapkan protokol keselamatan standar untuk ruang rontgen [10]. Evaluasi paparan dosis radiasi internal dan eksternal sangat penting mengingat dampak serius yang dialami pekerja radiasi dan lingkungan sekitar akibat paparan radiasi *X-Ray*. Selain dampak stokastik yang akan dialami manusia, dampak yang lebih serius juga akan terjadi pada lingkungan sekitar, misalnya pada ekosistem sekitar [11]. Telah dilakukan penelitian untuk mengevaluasi paparan radiasi internal terhadap radiasi *X-Ray* yang bersumber dari *CT-Scan* yang meningkatkan peningkatan proses metabolisme tubuh secara signifikan sehingga dapat menyebabkan kanker terutama pada anak-anak [12]. Penelitian mengukur dosis radiasi *X-Ray* menggunakan detektor radiasi dengan menilai sumber dan efeknya menyebutkan adanya ketidaksesuaian antara dosis serapan yang diperoleh dengan jaringan tubuh sebagai objek (target). Hal ini disebabkan oleh ketidakakuratan dalam proses penilaian dan hasilnya disebut kesalahan estimasi dosis [13]. Selain pemantauan radiasi secara berkala terhadap pekerja radiasi, tingkat keakuratan hasil paparan juga

dapat diukur menggunakan *survey* meter yang terkalibrasi dengan baik. Bahan yang digunakan pada *survey* meter juga mempengaruhi keakuratan data yang diperoleh, namun hal ini dapat diminimalisir dengan mengetahui data yang diperoleh dari hasil paparan menggunakan *machine learning* [14].

Penelitian yang dilakukan untuk mengetahui tingkat keamanan paparan radiasi X-Ray masih menggunakan cara konvensional yaitu secara manual dengan pemantauan berkala menggunakan sistem dosimetri. Pengumpulan data dilakukan untuk memperoleh hasil yang aman dan tidak aman berdasarkan analisis selanjutnya. Hal ini sangat merepotkan, dan tingkat efektivitasnya perlu ditingkatkan [15]. Teknik lain untuk mengidentifikasi keakuratan hasil deteksi paparan radiasi adalah dengan menggunakan *Convolutional Neural Network* (CNN) yang memerlukan pengambilan gambar X-Ray dengan resolusi yang baik. Proses ini menyulitkan operator dalam memilih data awal yaitu klasifikasi komposisi ketajaman gambar [16].

**Tabel 1.** Tabel perbandingan teknik identifikasi data dosis radiasi

| Kumpulan Data REF |                     |       | Teknik Identifikasi  | Membatasi  |
|-------------------|---------------------|-------|--|--|
| [6]               | Nilai radiasi       | dosis | Simulasi Monte Carlo   | <ul style="list-style-type: none"> <li>- Alat ukur yang mahal</li> <li>- Perlu beberapa metode lanjutan untuk mendapatkan hasil akhir (analisis data)</li> </ul>                                   |
| [14]              | Nilai radiasi       | dosis | Secara manual, nilai dosis paparan harian dikumpulkan menggunakan TLD kemudian dirata-ratakan setiap tahunnya menggunakan <i>Landauer Database</i> | <ul style="list-style-type: none"> <li>- Membatasi kegunaan <i>database</i> secara keseluruhan masa hidup.</li> <li>- Perlu lebih banyak penyimpanan</li> <li>- Radiasi hamburan tinggi</li> </ul> |
| [15]              | Gambar X-Ray        |       | Perhitungan Dosis PCXMC Fitur fusi   | <ul style="list-style-type: none"> <li>- Pekerja tetap diperbolehkan mendapatkan dosis berlebih</li> <li>- Perlu modalitas tinggi untuk instal perangkat lunak</li> </ul>                          |
| [16]              | Gambar rontgen dada |       | HOG<br>CNN   | <ul style="list-style-type: none"> <li>- Perlu modalitas tinggi untuk instal perangkat lunak</li> <li>- Perlu beberapa metode lanjutan sebelum gambar X-Ray dapat diidentifikasi</li> </ul>        |

Penelitian ini bertujuan untuk mengidentifikasi data dosis paparan radiasi X-Ray menggunakan *Random Forest Classification*, yang dibandingkan dengan *Naïve Bayes* dan *K Nearest Neighbor* (KNN) untuk keakuratan data. Dengan cara ini akan diperoleh identifikasi tingkat keamanan radiasi yang paling akurat. Penelitian ini bertujuan untuk membangun sistem identifikasi data dosis paparan radiasi X-Ray dengan akurasi tinggi. Jadi, meningkatkan kinerja dan akurasi dari data pengukuran paparan radiasi X-Ray dalam kalibrasi atau *Quality Assurance* (QA).

Tabel 1 menunjukkan bahwa sebagian besar dari mereka mempunyai teknik identifikasi yang perlu lebih banyak dan tingkat akurasinya rendah. Oleh karena itu diperlukan suatu sistem untuk mengidentifikasi data tersebut agar hasil paparan radiasi X-Ray lebih akurat dan efektif. Penelitian ini membangun sistem untuk mengidentifikasi

tingkat keamanan klasifikasi dosis radiasi X-Ray menggunakan *Random Forests*. Hal ini diberikan agar efektif dan mendapatkan akurasi yang tinggi. Proses klasifikasi pelaporan data dosis radiasi X-Ray bisa lebih efisien karena tidak lagi digunakan secara manual. Keakuratan data diperoleh dengan melakukan pemeriksaan menggunakan radiasi pengion untuk meminimalkan dampak buruk jangka panjang. Pemilihan faktor paparan dapat dilakukan secara optimal sehingga dosis paparan radiasi yang diterima pekerja radiasi dan lingkungan sekitar mudah dikendalikan seminimal mungkin.

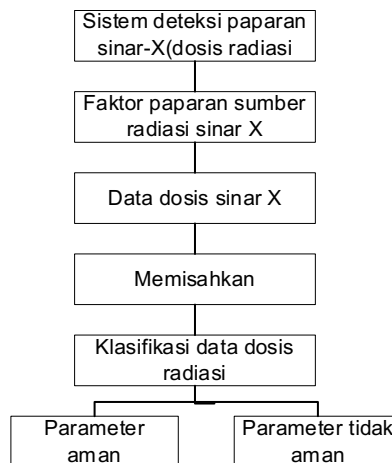
## 2. METODE

Data yang digunakan dalam penelitian ini diperoleh dari pengujian dosis radiasi X-Ray berdasarkan parameter potensial tabung (kV), arus tabung (mA), dan waktu pemaparan (s) sumber X-Ray dalam satuan Sv/h menggunakan meteran survei digital. Dengan mengetahui nilai kuantitatif pada dosis ekuivalen maka tingkat keamanan hasil paparan radiasi dari aktivitas. Cara ini dapat membantu pekerja radiasi, lingkungan sekitar sumber radiasi, dalam melaksanakan kegiatan proteksi radiasi. Oleh karena itu, data yang akurat diperoleh dengan melakukan pemeriksaan menggunakan radiasi pengion untuk meminimalkan efek buruk jangka panjang. Selain itu diperoleh parameter penentuan potensial tabung optimal (kV), arus tabung (mA), dan waktu pemaparan (s). sumber X-Ray tertentu juga akan diketahui. Proses identifikasi data dosis radiasi ditunjukkan pada Gambar 1.

Metode yang digunakan untuk mengidentifikasi dosis paparan radiasi pada sumber X-Ray terdiri dari pembuatan sistem pendeteksi radiasi X-Ray dan melakukan pemaparan oleh sumber X-Ray dengan berbagai parameter tegangan dan arus sumber.

Variasi tersebut bertujuan untuk menentukan tingkat perlindungan yang harus dimiliki masyarakat di sekitar sumber X-Ray. Parameter ini ditetapkan untuk mengetahui tingkat perlindungan yang harus dimiliki orang-orang di sekitar sumber X-Ray agar diperoleh parameter yang paling optimal [17]. Setelah data dari variasi faktor paparan dikumpulkan, pencarian titik aman dan tidak aman pun dilakukan. Nilai batas dosis ekuivalen diperoleh dari banyaknya dosis yang diserap oleh organ di sekitarnya dengan menggunakan alat detektor [18]. Identifikasi dosis radiasi X-Ray dapat dilakukan dengan menggunakan berbagai metode kecerdasan buatan. Data dosis radiasi yang diperoleh melalui sistem (dalam  $\mu\text{Sv/h}$ ) kemudian dipisahkan menjadi dua bagian yaitu aman dan tidak aman.

Selain hanya menggunakan satu metode, perbandingan juga dapat dilakukan dengan menggunakan beberapa metode untuk mendapatkan akurasi yang tinggi. Pada penelitian ini dilakukan perbandingan terhadap tiga hasil pencarian untuk mendapatkan tingkat akurasi yang berbeda.



**Gambar 1.** Proses identifikasi dosis radiasi

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Faktor Paparan

Dosis radiasi yang terpapar pada pekerja harus dijaga serendah mungkin (ALARA), dan salah satu cara untuk memastikannya adalah dengan memberikan parameter faktor paparan yang optimal. Pada X-Ray bergerak, dosis yang diperoleh jauh lebih tinggi karena penggunaan parameter arus tabung yang tinggi dan jarak pemeriksaan yang pendek.

Pada penelitian sebelumnya, dosis paparan radiasi dipengaruhi oleh arus tabung dan jarak paparan dari sumber [19]. Pemberian waktu paparan yang lama juga mempengaruhi dosis paparan radiasi yang diterima pekerja dan lingkungan sekitar. Semakin lama waktu pemaparan maka semakin tinggi pula dosis yang diterima [20], [21].

Faktor paparan yang digunakan pada penelitian ini adalah variasi potensial tabung X-Ray 50 kV, 60 kV, 70 kV dan 80 kV. Untuk setiap variasi potensial tabung juga diberikan variasi arus tabung sebesar 16 mA, 32 mA, 63 mA dan 100 mA serta waktu pemaparan yang berbeda sebesar 0,1 s; 0,16 detik; 0,2 detik; 0,32 detik; 0,4 detik; 0,5 detik; 0,63 detik; 0,8 detik; 1,0 detik dan 2,0 detik disetel. Variasi parameter yang digunakan disesuaikan dengan spesifikasi *mobile X-Ray* (MEDNIF tipe SF-100BY) dengan jarak ke sumber X-Ray 50 meter. Secara matematis hubungan antara faktor paparan dengan dosis paparan radiasi yang diterima diberikan oleh Persamaan (1) [21], [22]:

$$D(mR) = \frac{P \text{ mAs } (kV)^2}{r^2} \quad (1)$$

Keterangan:

D : Dosis paparan dalam mR

P : Faktor mesin X-Ray ( $P = 15$ )

mA : arus tabung

kV : potensial tabung

R : jarak dari sumber (dalam cm)

konversi dosis paparan dari mR ke  $\mu\text{Sv}$  adalah 1 mR sama dengan 87  $\mu\text{Sv}$ .

### 3.2. Proses Identifikasi

Dalam penelitian ini menggunakan validasi silang dengan lipatan = 5. Validasi silang merupakan prosedur yang paling banyak digunakan untuk melakukan evaluasi kinerja prediktif dari suatu model. Validasi silang dapat memperkirakan hingga generalisasi dalam suatu metode pemodelan [23]. Jumlah semua data akan terpisah menjadi lima bagian yang sama.

Pelajaran ini menggunakan 160 data set dari data dosis paparan radiasi X-Ray, sehingga masing-masing bagian terdiri dari 32 data secara acak. Jika bagian kelima adalah A, B, C, D, dan E, maka pengujian data dilakukan sesuai Tabel 2 [24]. Pada Tabel 2 terdapat 5 langkah validasi silang data. Langkahnya sebagai berikut:

*Split* 1: pelatihan menggunakan A, B, C, D dan pengujian menggunakan E, diperoleh ketelitian = a

*Split* 2: pelatihan menggunakan A, B, C, E dan pengujian menggunakan D, diperoleh ketelitian = b

*Split* 3: pelatihan menggunakan A, B, D, E dan pengujian menggunakan C, diperoleh ketelitian = c

*Split* 4: pelatihan menggunakan A, C, D, E dan pengujian menggunakan B, diperoleh ketelitian = d

*Split* 5: pelatihan menggunakan B, C, D, E dan pengujian menggunakan A, diperoleh ketelitian = e

Jadi rata-rata keakuratan data dapat dilihat pada Persamaan (2) [25]:

$$\frac{a+b+c+d+e}{5} \quad (2)$$

Data hasil pemodelan menggunakan *cross validation* kemudian dilakukan identifikasi menggunakan *Random Forest Classification*, *Naïve Bayes*, dan KNN. Metode *Random Forest Classification* digunakan untuk membangun pohon keputusan dengan mengambil data secara acak [26].

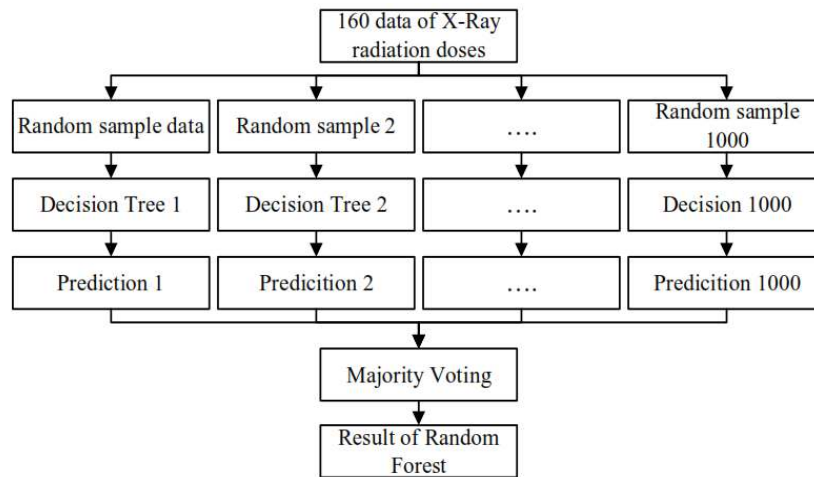
**Tabel 2.** Proses data validasi silang

|            | A       | B       | C       | D       | E       |
|------------|---------|---------|---------|---------|---------|
| Pisahkan 1 | 32 data | 32 data | 32 data | 32 data | 32 data |
| Pisahkan 2 | 32 data | 32 data | 32 data | 32 data | 32 data |
| Pisahkan 3 | 32 data | 32 data | 32 data | 32 data | 32 data |
| Pisahkan 4 | 32 data | 32 data | 32 data | 32 data | 32 data |
| Pisahkan 5 | 32 data | 32 data | 32 data | 32 data | 32 data |

Dengan menggunakan beberapa variasi parameter faktor paparan, diperoleh 160 data dengan 87 milik. Data dosis aman (nilai  $\dot{y}$  5,707  $\mu\text{Sv/h}$ ) dan sisanya 73 data dosis tidak aman (nilai  $\dot{y}$  5,707  $\mu\text{Sv/h}$ ). Data tersebut kemudian diidentifikasi menggunakan metode *Random Forest Classification*. Data dipisahkan menjadi dua kategori: data uji dan data pelatihan. Total 160 *data set* dibagi menjadi 75% data pengujian dan 25% data pelatihan (masing-masing 120 dan 40 sampel). Proses ini menghasilkan kumpulan sampel pelatihan. Himpunan ini kemudian diuji menggunakan data pengujian acak dalam sebuah matriks. Hal ini, pada gilirannya, menghasilkan data pengujian yang dikenal sebagai pohon klasifikasi. Pada penelitian ini terdapat 100 pohon pengklasifikasi yang

menghasilkan nilai *Random Forest*. Penjumlahan beberapa pohon klasifikasi memberikan nilai akurasi, sensitivitas (TPR), *spesifisitas* (TNR) dan *False Positive Rate* (FPR) [27].

Proses klasifikasi di *Random Forest* dimulai dengan membagi 160 data sampel (87 data aman dan 73 data tidak aman) menjadi 1000 sampel acak. Dari setiap sampel acak diperoleh 1000 pohon keputusan dan kemudian diperoleh 1000 data prediksi. Setelah data prediksi terbentuk kemudian dilakukan pemungutan suara prediksi ke setiap kelas. Selanjutnya, gabungkan hasil suara dari setiap kelas dan dapatkan suara mayoritas untuk semua. Pemungutan suara mayoritas adalah yang terbanyak hasil prediksi. Proses klasifikasi data *Random Forest* pada penelitian ini, dapat ditunjukkan pada Gambar 2.



**Gambar 2.** Proses klasifikasi data *random forest*

Beberapa istilah yang digunakan dalam pengujian hasil data dosis paparan radiasi antara lain *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Istilah ini digunakan dalam akurasi, sensitivitas (TPR), *spesifisitas* (TNR) dan angka positif palsu (FPR). TP merupakan kondisi aktual (aktual) dan hasil prediksi untuk mengidentifikasi dosis paparan radiasi yang “aman”. TN diperoleh kondisi aktual (aktual) dan hasil prediksi untuk mengidentifikasi dosis paparan radiasi yang “tidak aman”. FP diperoleh jika kondisi “tidak aman” namun hasil deteksi “aman”, sedangkan nilai FN diperoleh bila kondisi aktual (aktual) “aman” namun diketahui hasil (prediksi) “tidak aman” [19], [21], [28]. Akurasi diperoleh dengan menggunakan Persamaan (3) [29]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad 0,0 < Accuracy < 1,0 \quad (3)$$

Selain itu, pengujian juga digunakan untuk menentukan sensitivitas, yaitu matriks standar untuk mengidentifikasi jumlah data dosis radiasi dalam kumpulan data yang tidak seimbang. Sensitivitas dihitung menggunakan Persamaan (4) [30]:

$$Sensitivity = \frac{TP}{TP+FN} \quad (4)$$

Validitas data yang diidentifikasi dapat ditentukan dengan menggunakan *spesifisitas* (TNR) seperti yang diberikan pada Persamaan (5). Kekhususan mengacu pada

kuantifikasi kemampuan metode dalam mengklasifikasikan data dengan benar sehingga memungkinkan memperoleh hasil yang sesuai [31].

$$Spesificity = \frac{TN}{TN+FP} \quad (5)$$

*Fall out* atau *false positive rate* dapat ditentukan dengan menggunakan rasio FP dibagi jumlah FP dan TN, seperti yang diberikan pada Persamaan (6) [32]:

$$False\ Positive\ Rate = \frac{FP}{FP+TN} \quad (6)$$

Dan kemudian menentukan hasil *Confusion Matrix*. Gambar 3 merupakan *Confusion Matrix* secara umum berdasarkan pengelompokan data dosis paparan radiasi X-Ray yang aman dan tidak aman. *Confusion Matrix* berfungsi untuk memetakan pengelompokan antara data aktual dan data prediksi [33]. Berdasarkan hasil *Confusion Matrix* diperoleh kurva *Receiver Operation Characteristic* (ROC).

|            |        | Actual         |                |
|------------|--------|----------------|----------------|
|            |        | Safe           | Unsafe         |
| Prediction | Safe   | True Positive  | False Negative |
|            | Unsafe | False Positive | True Negative  |

**Gambar 3.** *Confusion matrix* secara umum

### 3.3. Analisis Kinerja

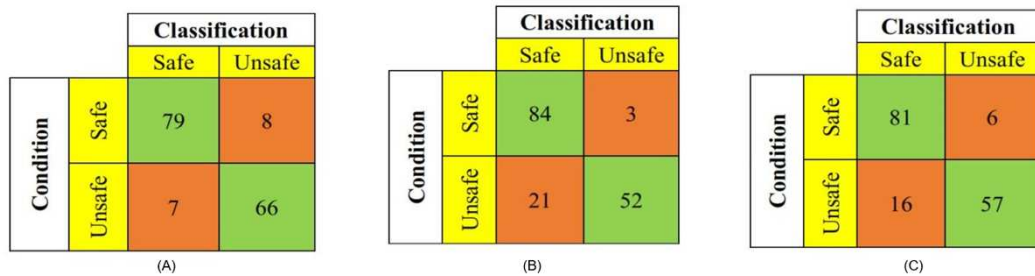
Setelah mendapatkan data dari *Confusion Matrix*, buatlah kurva karakteristik operasi penerima (ROC). Hubungan antara skor tingkat FP dan tingkat TP mewakili kurva ini. Validasi hasil pengukuran menggunakan kurva ROC dan *Confusion Matrix* untuk mendapatkan skor tertinggi. Kurva ROC menggambarkan kinerja klasifikasi tanpa distribusi kelas atau kesalahan [34]. Daerah di bawah kurva disebut *Area Under Curve* (AUC) berdasarkan bentuk kurvanya.

*Under field* memiliki skor antara 0% dan 100%. Namun yang menarik untuk dihitung adalah skornya di atas 50%. Area bagian bawahnya tinggi, sehingga akan membaik [35]. Interpretasi nilai *Area Under Receiver Operating Characteristic* (AUC) adalah diberikan pada Tabel 3 [36].



**Tabel 3.** Proses data validasi silang

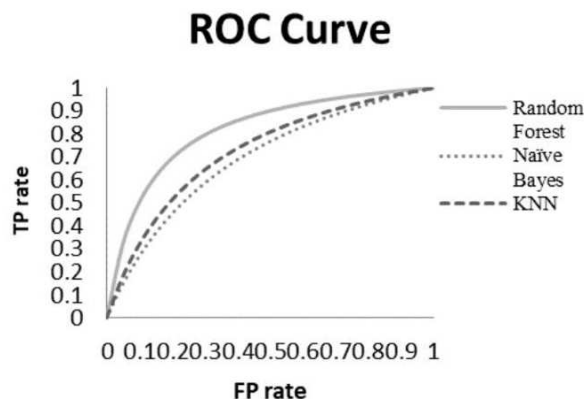
| Rentang           | AUC          |
|-------------------|--------------|
| >50,0 % - 60,0 %  | Sangat lemah |
| >60,0 % - 70,0 %  | Lemah        |
| >70,0 % - 80,0 %  | Adil         |
| >80,0 % - 90,0 %  | Bagus        |
| >90,0 % - 100,0 % | Sangat bagus |

**Gambar 4.** Kurva ROC dari ketiga metode

Hasil *Confusion Matrix* data dosis paparan radiasi X-Ray menggunakan *Random Forest Classification* dapat dilihat pada Gambar 4 (a). Berdasarkan keseluruhan data berjumlah 160 data, 87 data dosis aman dan 73 data dosis tidak aman. Setelah dilakukan perhitungan data, berdasarkan *Confusion Matrix* 2x2 diperoleh TP sebanyak 79 data, FN sebanyak 8 data, FP sebanyak 7 data, dan TN sebanyak 66 data.

Hasil *Confusion Matrix* data dosis paparan radiasi X-Ray menggunakan *Naïve Bayes* dapat ditunjukkan pada Gambar 4 (b). Berdasarkan keseluruhan data berjumlah 160 data, 87 data dosis aman dan 73 data dosis tidak aman. Setelah dilakukan perhitungan data, berdasarkan *Confusion Matrix* 2x2 diperoleh TP sebanyak 84 data, FN sebanyak 3 data, FP sebanyak 21 data, dan TN sebanyak 52 data.

Hasil *Confusion Matrix* data dosis paparan radiasi X-Ray menggunakan KNN dapat ditampilkan pada 160 data, 87 data dosis aman dan 73 data dosis tidak aman. Setelah dilakukan perhitungan data, berdasarkan *Confusion Matrix* 2x2 diperoleh TP sebanyak 81 data, FN sebanyak 6 data, FP sebanyak 16 data, dan TN sebanyak 57 data. Hasil dari *Random Forest Classification* telah dibandingkan dengan *Naïve Bayes* dan KNN. Tabel 3 merupakan hasil perbandingan dari ketiga metode.

**Gambar 5.** Kurva hubungan antara FP rate dan TP rate

### 3.4. Hasil Analisis Kinerja

Hubungan nilai sensitivitas dan *spesifisitas* yang diperoleh kurva ROC menghasilkan nilai AUC. Gambar 5 merupakan kurva ROC, FP rate merupakan nilai *spesifisitas* dan TP rate merupakan nilai sensitivitas. Itu menunjukkan kinerja metode yang digunakan. Analisis kinerja suatu metode *machine learning* dikatakan lebih baik jika nilai AUC mendekati 100% [28], [39].

Berdasarkan kurva ROC pada Gambar 5, pada titik *False Positive Rate* yang sama, kurva *Random Forest* berada pada titik *True Positive Rate* yang paling tinggi dibandingkan dengan kurva *Naïve Bayes* dan KNN. Dengan demikian, metode *Random Forest* dapat bekerja lebih baik dalam mengklasifikasikan fitur. Hasil tersebut juga diperkuat dengan menghitung nilai akurasi menggunakan Persamaan 3.

Dengan hasil akurasi tertinggi adalah metode *random forest* dan dapat dilihat pada Tabel 4

**Tabel 4.** Proses data validasi silang

| Teknik Identifikasi              | Sensitivitas | <i>Spesifisitas</i> | Analisis Kinerja | Akurasi | Tingkat Positif Palsu |
|----------------------------------|--------------|---------------------|------------------|---------|-----------------------|
| Klasifikasi <i>Random Forest</i> | 91%          | 90%                 | 9%               | 90%     | 97%                   |
| <i>Bayes Naif</i>                | 93%          | 78%                 | 22%              | 86%     | 85%                   |
| KNN                              | 96%          | 71%                 | 29%              | 85%     | 94%                   |

Tabel 4 menunjukkan bahwa metode yang diusulkan memiliki kinerja lebih baik dalam mengidentifikasi kondisi aman dan tidak aman. Metode yang diusulkan mencapai klasifikasi tertinggi dalam *spesifisitas*, akurasi dan analisis kinerja dibandingkan metode lain dalam kondisi dan data yang sama. Selain itu, metode yang diusulkan juga memiliki tingkat positif palsu yang lebih rendah, yang menunjukkan bahwa metode yang diusulkan memiliki kemungkinan lebih rendah dalam mengidentifikasi kasus negatif sebagai kasus positif dalam data. Berdasarkan hasil tersebut, metode *Random Forest Classification* dapat bekerja lebih baik dalam mengidentifikasi paparan radiasi X-Ray dibandingkan metode lainnya. Namun, metode yang diusulkan memiliki klasifikasi terendah dalam mengidentifikasi kasus positif dengan benar dibandingkan metode lainnya

## 4. KESIMPULAN

Mengukur tingkat paparan radiasi X-Ray sangatlah penting. Radiasi X-Ray dapat merusak struktur molekul. Oleh karena itu, pengukuran tingkat paparan radiasi X-Ray diperlukan untuk mencegah paparan berlebih pada tubuh manusia. Klasifikasi hutan secara acak diusulkan untuk mengukur dan menentukan tingkat keterpaparan dalam penelitian ini. Hasilnya menunjukkan bahwa metode yang diusulkan memiliki kinerja lebih baik dibandingkan metode lainnya, dengan *spesifisitas* 90%, akurasi 90%, dan analisis kinerja 97%. Selain itu, hasil yang diusulkan memiliki tingkat positif palsu yang lebih rendah yaitu sebesar 9%, yang menunjukkan kemungkinan lebih rendah untuk mengidentifikasi kasus negatif sebagai kasus positif dalam data. Di masa depan, kami menyarankan untuk menambah jumlah kumpulan data dan mengembangkan metode klasifikasi untuk meningkatkan kinerja.

## DAFTAR PUSTAKA

- [1] M. Zdora, "State of the Art of X-Ray Speckle-Based Phase-Contrast and Dark-Field Imaging," *J. Imaging*, vol. 4, no. 5, pp. 1–36, 2018, doi: 10.3390/jimaging4050060.
- [2] P. D. Quinn, M. Gomez-gonzalez, F. Cacho-nerin, and J. E. Parker, "Beam and sample movement compensation for robust spectro-microscopy measurements on a hard X-Ray nanoprobe," *J. Synchrotron Radiat.*, vol. 28, no. 5, pp. 1528–1534, 2021, doi: 10.1107/S1600577521007736.
- [3] E. E. Altunsoy, H. O. Tekin, A. Mesbahi, and I. Akkurt, "MCNPX Simulation for Radiation Dose Absorption of Anatomical Regions and Some Organs," *Acta Phys. Pol. A*, vol. 137, no. 4, pp. 561–565, 2020, doi: 10.12693/APhysPolA.137.561.
- [4] P. J. Withers *et al.*, "X-Ray computed tomography," *Nat. Rev. Methods Prim.*, vol. 1, no. 18, pp. 1–21, 2021, doi: 10.1038/s43586-021-00015-4.
- [5] F. T. M. Hounsou *et al.*, "Radiological Quality and Dangerousness of Ferrous and Non-ferrous Metals Waste in Cotonou (Benin)," *Asian J. Environ. Ecol.*, vol. 12, no. 2, pp. 13–21, 2020, doi: 10.9734/AJEE/2020/v12i230154.
- [6] Y. I. Cho, S. H. Shin, and J. H. Kim, "Dose assessment and shielding analysis of the working environments for non-destructive testing," *J. Instrum.*, vol. 15, no. June, pp. 1–11, 2020, doi: 10.1088/1748-0221/15/06/P06012.
- [7] M. Bonczyk, A. Grygier, and K. Skubacz, "'Quantum Pendants' - the measurement of exposure to enhanced natural radioactivity," *Measurement*, vol. 196, no. February, p. 111212, 2022, doi: 10.1016/j.measurement.2022.111212.
- [8] D. R. Ningtias, B. Wahyudi, and I. T. Harsoyo, "Comparative Test of the Effect of X-Ray Tube Current Analysis and Exposure Time on CR (Computed Radiography) Image Quality," *JITE (Journal Informatics Telecommun. Eng.)*, vol. 6, no. 1, pp. 267–275, 2022, doi: 10.31289/jite.v6i1.7334 Received:
- [9] S. Zhang *et al.*, "First measurements of the radiation dose on the lunar surface," *Sci. Adv.*, vol. 6, no. 39, pp. 1–6, 2020, doi: 10.1126/sciadv.aaz1334.
- [10] P. Pelliccioli *et al.*, "Study of the X-Ray radiation interaction with a multislit collimator for the creation of microbeams in radiation therapy," *J. Synchrotron Radiat.*, vol. 28, no. 2, pp. 392–403, 2021, doi: 10.1107/S1600577520016811.
- [11] T. Kang, W. Park, Y. Han, K. M. Bong, and K. Kim, "Natural and artificial radioactivity in volcanic ash soils of Jeju Island, Republic of Korea, and assessment of the radiation hazards: importance of soil properties," *J. Radioanal. Nucl. Chem.*, vol. 323, no. March, pp. 1113–1124, 2020, doi: 10.1007/s10967-020-07024-9.
- [12] I. A. Tsalafoutas, M. H. Kharita, H. Al-Naemi, and M. K. Kalra, "Radiation dose monitoring in computed tomography: Status, options and limitations," *Phys. Medica*, vol. 79, no. November, pp. 1–15, 2020, doi: 10.1016/j.ejmp.2020.08.020.
- [13] R. D. Daniels, G. M. Kendall, I. Thierry-chef, M. S. Linet, and H. M. Cullings, "Strengths and Weaknesses of Dosimetry Used in Studies of Low-Dose Radiation Exposure and Cancer," *JNCI Monogr.*, vol. 2020, no. 56, pp. 114–132, 2020, doi: 10.1093/jncimonographs/lgaa001.
- [14] R. C. Yoder *et al.*, "Dosimetry for the study of medical radiation workers with a focus on the mean absorbed dose to the lung, brain and other organs," *Int. J. Radiat. Biol.*, vol. 98, no. 4, pp. 619–630, 2018, doi: 10.1080/09553002.2018.1549756.
- [15] Z. Brady *et al.*, "Technique, radiation safety and image quality for chest X - ray imaging through glass and in mobile settings during the COVID - 19 pandemic," *Phys. Eng. Sci. Med.*, vol. 43, no. 3, pp. 765–779, 2020, doi: 10.1007/s13246-020-00899-8.
- [16] N.- A-Alam, M. Ahsan, M. A. Based, J. Haider, and M. Kowalski, "COVID-19 Detection from Chest X-Ray Images Using Feature Fusion and Deep Learning," *Sensors*, vol. 21, no. 4, p. 1480, 2021, doi: <https://doi.org/10.3390/s21041480>.

- [17] D. Rottke, L. Gohlke, R. Schrödel, S. Hassfeld, and D. Schulze, "Operator Safety During the Acquisition of Intraoral Images With A Handheld and Portable X-Ray Device," *Dentomaxillofac Radiol*, vol. 47, no. 3, pp. 1–9, 2018, doi: 10.1259/dmfr.20160410.
- [18] C. Clement *et al.*, "Keeping the ICRP recommendations fit for purpose," *J. Radiol. Prot.*, vol. 41, no. 4, pp. 1390–1409, 2021, doi: 10.1088/1361-6498/ac1611.
- [19] A. Lahham, A. Issa, and H. ALMasri, "Patient Radiation Dose from Chest X-Ray Examinations In The West Bank — Palestine," *Radiat. Prot. Dosimetry*, vol. 178, no. 3, pp. 298–303, 2017, doi: 10.1093/rpd/ncx110.
- [20] Y. F. Ali, F. A. Cucinotta, L. Ning-Ang, and G. Zhou, "Cancer Risk of Low Dose Ionizing Radiation," *Front. Phys.*, vol. 8, no. August, pp. 1–9, 2020, doi: 10.3389/fphy.2020.00234.
- [21] M. W. Alhamd and N. A. Abbood, "The effect voltage Device of X-Ray on radiation doses and image quality," *J. Phys. Conf. Ser. Pap.*, vol. 1664, no. 2020, pp. 1–12, 2020, doi: 10.1088/1742-6596/1664/1/012002.
- [22] C. R. T. Simo, O. N. Samba, P. K. Talla, and L. C. Fai, "Radiation dose from three-phase X-Ray machines: A comparison between different models," *Int. J. Radiat. Res.*, vol. 19, no. 3, pp. 559–567, 2021, doi: 10.18869/acadpub.ijrr.19.3.559.
- [23] B. I. J. Lamarca and S. C. Ambat, "The Development of a Performance Appraisal System Using Decision Tree Analysis and Fuzzy Logic," *Int. J. Intell. Eng. Syst.*, vol. 11, no. 4, pp. 11–19, 2017, doi: 10.22266/ijies2018.0831.02.
- [24] D. Berrar and M. Carlo, "Cross-Validation," in *Encyclopedia of Bioinformatics and Computational Biology*, S. Ranganathan, M. Gribskov, K. Nakai, C. Schönbach, Encyclopedia of Bioinformatics and Computational Biology, and Academic Press, Eds. Selangor: Elsevier Inc, 2019, pp. 542–545.
- [25] M. A. S. Ali and M. Abd-Elfattah, "A Hybrid Model ( SVM-LOA ) for Epileptic Seizure Detection in Long-Term EEG Records Using Machine Learning Techniques," *Int. J. Intell. Eng. Syst.*, vol. 11, no. 5, pp. 162–172, 2018, doi: 10.22266/ijies2018.1031.15.
- [26] U. L. Yuhana, V. R. Oktavia, C. Fatichah, and A. Purwarianti, "Automatic Assessment of Answers to Mathematics Stories Question Based on Tree Matching and Random Forest," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 2, pp. 200–212, 2022, doi: 10.22266/ijies2022.0430.19.
- [27] A. K. Nugroho, D. M. K. Nugraheni, T. A. Putranto, I. K. E. Purnama, and M. H. Purnomo, "Utilizing the Hepta Convolutional Layer Neural Network ( HCL-NN ) Based on a Multi Optimizer for the Classification of Brain Stroke MRI," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 3, pp. 304–318, 2022, doi: 10.22266/ijies2022.0630.26.
- [28] F. Itoo, Meenakshi, and S. Singh, "Comparison and analysis of logistic regression, Nai"ve Bayes and KNN machine learning algorithms for credit card fraud detection," *Int. J. Inf. Technol.*, vol. 13, no. February, pp. 1503–1511, 2020, doi: 10.1007/s41870-020-00430-y.
- [29] C. Iwendi *et al.*, "COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm," *Front. Public Heal.*, vol. 8, no. July, pp. 1–9, 2020, doi: 10.3389/fpubh.2020.00357.
- [30] V. G. Rangappa, "Classification of Cardiac Arrhythmia stages using Hybrid Features Extraction with K-Nearest Neighbour classifier of ECG Signals," *Int. J. Intell. Eng. Syst.*, vol. 11, no. 6, pp. 21–32, 2018, doi: 10.22266/ijies2018.1231.03.
- [31] P. Coelho, A. Pereira, and M. Salgado, "A Deep Learning Approach for Red Endoscopies," in *International Conference Image Analysis and Recognition*, 2018, vol. 10882, pp. 553–561, doi: 10.1007/978-3-319-93000-8\_63.
- [32] T. Akter *et al.*, "Improved Transfer-Learning-Based Facial Recognition Framework to Detect Autistic Children at an Early Stage," *brain Sci.*, vol. 11, no. 6, p. 734, 2021, doi: <https://doi.org/10.3390/brainsci11060734>.

- [33] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," *IEEE Access*, vol. 8, pp. 90847–90861, 2020, doi: 10.1109/ACCESS.2020.2994222.
- [34] E. T. Lau, L. Sun, and Q. Yang, "Modelling , prediction and classification of student academic performance using artificial neural networks," *SN Appl. Sci.*, vol. 1, no. 9, pp. 1–10, 2019, doi: 10.1007/s42452-019-0884-7.
- [35] S. A. Ali *et al.*, "GIS-based landslide susceptibility modeling: A comparison between fuzzy multi-criteria and machine learning algorithms," *Geosci. Front.*, vol. 12, no. 2, pp. 857–876, 2021, doi: 10.1016/j.gsf.2020.09.004.
- [36] J. Muschelli, "ROC and AUC with a Binary Predictor: a Potentially Misleading Metric," *J. Classif.*, vol. 37, no. October 2020, pp. 696–708, 2019, doi: <https://doi.org/10.1007/s00357-019-09345-1>.
- [37] P. A. Oakley, N. N. Ehsani, and D. E. Harrison, "The Scoliosis Quandary : Are Radiation Exposures From Repeated X-Rays Harmful ?," *Dose Response*, vol. 17, no. 2, pp. 1–10, 2019, doi: 10.1177/1559325819852810.
- [38] K. Ji *et al.*, "Research Progress on the Biological Effects of Low-Dose Radiation in China," *Dose Response*, vol. 26, no. 17, pp. 1–16, 2019, doi: 10.1177/1559325819833488.
- [39] H. Abbad Ur Rehman, C. Y. Lin, Z. Mushtaq, and S. F. Su, "Performance Analysis of Machine Learning Algorithms for Thyroid Disease," *Arab. J. Sci. Eng.*, vol. 46, no. 10, pp. 9437–9449, 2021, doi: 10.1007/s13369-020-05206-x.