
Ovarian Cancer Prediction Using Random Forest and Explainable Artificial Intelligence Analysis with SHAP

Hieronimus Feldi¹⁾, Wahyuni²⁾, dan Pitrasacha Adytia³⁾

^{1,2,3}Teknik Informatika, STMIK Widya Cipta Dharma^{1,2,3}

Jl. M Yamin No. 25, Gn.Kelua, Samarinda, Kalimantan Timur, 75123

E-mail: 2243060@wicida.ac.id¹⁾, wahyuni@wicida.ac.id²⁾, pitra@wicida.ac.id³⁾

ABSTRACT

Ovarian cancer has a high mortality rate due to delayed diagnosis, highlighting the need for an accurate and interpretable prediction model. This study aims to predict ovarian cancer using the Random Forest algorithm and analyze model interpretability through Explainable Artificial Intelligence (XAI) with SHapley Additive exPlanations (SHAP). The study employed the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology using an open medical dataset. Data preprocessing included cleaning, encoding, feature transformation, normalization, and class balancing using the Synthetic Minority Over-sampling Technique (SMOTE). Model performance was evaluated using a confusion matrix based on accuracy, precision, recall, and F1-score. The Random Forest model achieved an accuracy of 0.86, precision of 0.87, recall of 0.86, and F1-score of 0.86. Feature Importance and SHAP analyses identified Parity, miRNA, DNAMethylation, CA125, BMI, GeneExpression, and Age as the most influential features. The developed model was successfully implemented in a web-based application, demonstrating that the combination of Random Forest and SHAP provides accurate predictions with interpretable results to support ovarian cancer prediction.

Keywords: Ovarian Cancer, Random Forest, Explainable Artificial Intelligence, SHapley Additive exPlanations (SHAP), Machine Learning.

Prediksi Kanker Ovarium Menggunakan Random Forest dan Analisis Explainable AI dengan SHAP

ABSTRAK

Kanker ovarium memiliki tingkat mortalitas yang tinggi akibat keterlambatan diagnosis sehingga diperlukan model prediksi yang akurat dan mudah diinterpretasikan. Penelitian ini bertujuan memprediksi kanker ovarium menggunakan algoritma *Random Forest* serta menganalisis interpretabilitas hasil prediksi melalui pendekatan *Explainable Artificial Intelligence* (XAI) dengan metode *SHapley Additive exPlanations* (SHAP). Penelitian menggunakan metodologi *Cross Industry Standard Process for Data Mining* (CRISP-DM) dengan dataset medis terbuka. Tahap prapengolahan data meliputi *data cleaning*, *encoding*, transformasi fitur, normalisasi, dan penanganan ketidakseimbangan data menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE). Model dievaluasi menggunakan *confusion matrix* berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa model *Random Forest* memperoleh *accuracy* sebesar 0,86, *precision* sebesar 0,87, *recall* sebesar 0,86, dan *F1-score* sebesar 0,86. Analisis *Feature Importance* dan SHAP menunjukkan bahwa fitur Parity, miRNA, DNAMethylation, CA125, BMI, GeneExpression, dan Age merupakan faktor yang paling berpengaruh terhadap hasil prediksi. Model yang dikembangkan berhasil diimplementasikan pada aplikasi berbasis web serta mampu menghasilkan prediksi yang akurat dan mudah diinterpretasikan melalui penjelasan kontribusi setiap fitur menggunakan SHAP.

Kata Kunci: Kanker Ovarium, *Random Forest*, *Explainable Artificial Intelligence*, SHAP, *Machine Learning*.

1. PENDAHULUAN

Kanker ovarium merupakan salah satu jenis kanker ginekologi dengan tingkat mortalitas yang tinggi karena umumnya terdiagnosis pada stadium lanjut akibat gejala awal yang tidak spesifik. Keterlambatan diagnosis menyebabkan peluang keberhasilan pengobatan menjadi lebih rendah serta meningkatkan risiko kematian. Oleh karena itu, diperlukan suatu metode yang mampu membantu memprediksi risiko kanker ovarium secara

lebih akurat sehingga dapat mendukung pengambilan keputusan dalam bidang kesehatan.

Perkembangan *Machine Learning* telah dimanfaatkan dalam berbagai penelitian untuk membantu prediksi penyakit berdasarkan data medis. Salah satu algoritma yang banyak digunakan adalah *Random Forest* karena mampu menghasilkan performa klasifikasi yang baik, tahan terhadap *overfitting*, serta mampu mengolah data dengan jumlah fitur yang besar. Namun, hasil prediksi dari

model *Machine Learning* umumnya masih sulit dijelaskan karena bersifat *black box*, sehingga diperlukan pendekatan yang dapat meningkatkan transparansi model.

Beberapa penelitian terdahulu menunjukkan bahwa pendekatan *Explainable Artificial Intelligence* (XAI) menggunakan metode *SHapley Additive exPlanations* (SHAP) mampu menjelaskan kontribusi setiap fitur terhadap hasil prediksi model. Meskipun demikian, penelitian yang mengombinasikan algoritma *Random Forest* dengan metode SHAP untuk prediksi kanker ovarium masih terbatas. Oleh karena itu, diperlukan penelitian yang tidak hanya menghasilkan model prediksi dengan performa yang baik, tetapi juga mampu memberikan penjelasan terhadap hasil prediksi yang dihasilkan.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan membangun model prediksi kanker ovarium menggunakan algoritma *Random Forest* dan menganalisis hasil prediksi menggunakan pendekatan *Explainable Artificial Intelligence* (XAI) melalui metode *SHapley Additive exPlanations* (SHAP). Penelitian ini diharapkan menghasilkan model yang akurat, transparan, dan mudah diinterpretasikan sehingga dapat menjadi referensi dalam pengembangan sistem prediksi penyakit berbasis *machine learning*, khususnya pada kasus kanker ovarium.

2. RUANG LINGKUP

Penelitian ini mencakup analisis *Explainable Artificial Intelligence* (XAI) menggunakan metode *SHapley Additive exPlanations* (SHAP) untuk menjelaskan hasil prediksi kanker ovarium yang dihasilkan oleh algoritma *Random Forest*. Penelitian menggunakan dataset medis terbuka yang diperoleh dari Kaggle dengan fokus pada interpretasi kontribusi setiap fitur terhadap hasil prediksi. Batasan penelitian meliputi penggunaan algoritma *Random Forest* sebagai model klasifikasi dan SHAP sebagai metode interpretasi, sedangkan evaluasi model dilakukan menggunakan *Confusion Matrix* berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian berupa model prediksi kanker ovarium yang mampu memberikan hasil prediksi beserta penjelasan kontribusi setiap fitur menggunakan SHAP dan diimplementasikan dalam aplikasi berbasis web.

3. BAHAN DAN METODE

Main heading ditulis dengan huruf kapital semua. Penomoran *heading* dengan sistem *arabic* dengan *sub-heading* maksimal 3 tingkat.

3.1 Random Forest

Random Forest merupakan algoritma *machine learning* yang digunakan untuk melakukan klasifikasi dengan membangun sejumlah *decision tree* dari data pelatihan. Setiap *decision tree* menghasilkan prediksi yang kemudian digabungkan menggunakan mekanisme *majority voting* sehingga menghasilkan keputusan akhir. Algoritma ini dipilih karena memiliki

kemampuan klasifikasi yang baik, mampu mengurangi *overfitting*, serta efektif dalam mengolah data dengan jumlah fitur yang banyak.

3.2 Explainable Artificial Intelligence (XAI)

Explainable Artificial Intelligence (XAI) merupakan pendekatan yang bertujuan meningkatkan transparansi model *machine learning* sehingga hasil prediksi dapat dipahami oleh pengguna. Dalam penelitian ini, metode *SHapley Additive exPlanations* (SHAP) digunakan untuk menjelaskan kontribusi setiap fitur terhadap hasil prediksi kanker ovarium. Melalui SHAP, pengaruh masing-masing fitur terhadap peningkatan atau penurunan nilai prediksi dapat divisualisasikan sehingga proses pengambilan keputusan menjadi lebih mudah diinterpretasikan.

3.3 Python

Python merupakan bahasa pemrograman yang digunakan dalam penelitian ini untuk melakukan prapengolahan data, membangun model *Random Forest*, melakukan evaluasi model, serta mengimplementasikan analisis SHAP. Python dipilih karena memiliki sintaks yang sederhana dan didukung berbagai pustaka *machine learning* seperti *scikit-learn*, *SHAP*, *pandas*, dan *NumPy*, sehingga proses pengembangan model dapat dilakukan secara efektif.

3.4 Flask

Flask merupakan *framework* berbasis Python yang digunakan untuk mengembangkan aplikasi web pada penelitian ini. Flask digunakan untuk menghubungkan model prediksi *Random Forest* dengan antarmuka pengguna sehingga pengguna dapat memasukkan data, memperoleh hasil prediksi kanker ovarium, serta melihat penjelasan kontribusi setiap fitur berdasarkan analisis SHAP melalui halaman web.

3.5 Metode Penelitian

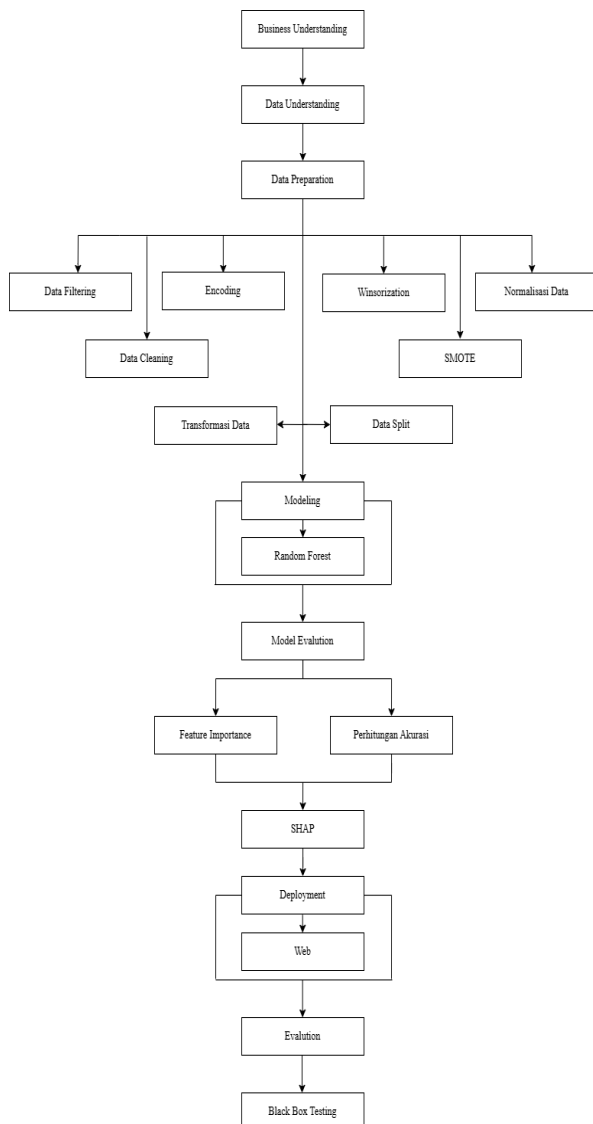
Penelitian ini menggunakan metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang terdiri atas enam tahapan, yaitu:

1. *Business Understanding*, yaitu mengidentifikasi permasalahan serta tujuan penelitian dalam membangun model prediksi kanker ovarium yang akurat dan mudah diinterpretasikan
2. *Data Understanding*, yaitu mengumpulkan dan menganalisis karakteristik dataset medis terbuka yang diperoleh dari Kaggle
3. *Data Preparation*, yaitu melakukan prapengolahan data yang meliputi *data cleaning*, *encoding*, transformasi data, *winsorization*, normalisasi, *data splitting*, serta penanganan ketidakseimbangan kelas menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE)
4. *Modeling*, yaitu membangun model prediksi menggunakan algoritma *Random Forest* berdasarkan data yang telah diproses
5. *Evaluation*, yaitu mengevaluasi performa model menggunakan *Confusion Matrix*

berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta melakukan analisis interpretabilitas menggunakan metode *SHapley Additive exPlanations* (SHAP)

6. *Deployment*, yaitu mengimplementasikan model ke dalam aplikasi berbasis web sehingga pengguna dapat melakukan prediksi kanker ovarium beserta melihat penjelasan kontribusi setiap fitur terhadap hasil prediksi.

Alur penelitian dimulai dari pengumpulan dataset medis terbuka, dilanjutkan dengan proses prapengolahan data, pembangunan model *Random Forest*, evaluasi performa model, analisis interpretabilitas menggunakan SHAP, dan implementasi model ke dalam aplikasi berbasis web. Tahapan tersebut menghasilkan sistem yang mampu memprediksi kanker ovarium sekaligus memberikan penjelasan mengenai kontribusi setiap fitur terhadap hasil prediksi. Alur tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 3.1. Alur Tahapan CRISP-DM

4. PEMBAHASAN

Penelitian ini menghasilkan model prediksi kanker ovarium menggunakan algoritma *Random Forest* yang dilengkapi dengan analisis *Explainable Artificial Intelligence* (XAI) menggunakan metode *SHapley Additive exPlanations* (SHAP). Model dibangun berdasarkan dataset medis terbuka yang telah melalui tahapan prapengolahan data, kemudian dievaluasi menggunakan *Confusion Matrix* berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, metode SHAP digunakan untuk menjelaskan kontribusi setiap fitur terhadap hasil prediksi sehingga meningkatkan transparansi dan kemudahan interpretasi model. Pada bagian ini dibahas hasil evaluasi model, analisis interpretabilitas menggunakan SHAP, serta implementasi model dalam aplikasi berbasis web.

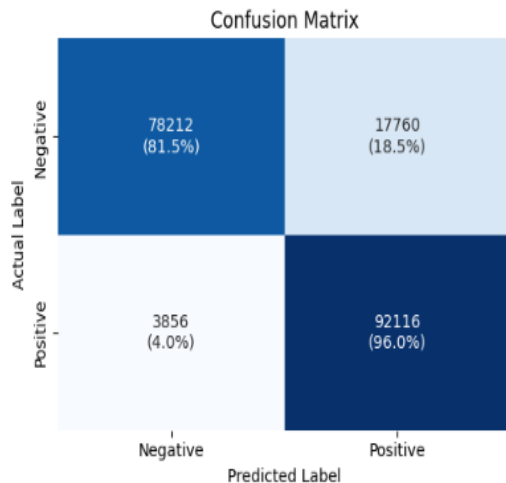
4.1 Hasil Prediksi Kanker Ovarium

Table 1. Hasil Evaluasi Model dengan SMOTE

Kelas	Precision	Recall	F1-Score
Negatif	0,93	0,78	0,84
Positif	0,81	0,94	0,87
Accuracy			0,86
Macro Avg	0,87	0,86	0,86
Weighted Avg	0,87	0,86	0,86

Berdasarkan Tabel 4.1, hasil evaluasi model *Random Forest* pada data latih yang telah diseimbangkan menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE) menunjukkan nilai *accuracy* sebesar 0,86, yang menunjukkan bahwa model mampu mengklasifikasikan risiko kanker ovarium dengan tingkat ketepatan sebesar 86%. Nilai *precision* sebesar 0,87 menunjukkan bahwa sebagian besar data yang diprediksi sebagai kelas positif merupakan prediksi yang benar, sedangkan nilai *recall* sebesar 0,86 menunjukkan bahwa model mampu mengenali sebagian besar data positif dengan baik. Nilai *F1-score* sebesar 0,86 menunjukkan keseimbangan yang baik antara *precision* dan *recall*, sehingga model memiliki performa klasifikasi yang baik.

Penggunaan metode SMOTE pada data latih bertujuan untuk mengatasi ketidakseimbangan distribusi kelas sehingga model dapat mempelajari karakteristik setiap kelas secara lebih seimbang. Dengan distribusi data yang lebih proporsional, algoritma *Random Forest* mampu menghasilkan performa klasifikasi yang baik tanpa cenderung berpihak pada kelas mayoritas. Selanjutnya, untuk mengetahui alasan di balik hasil prediksi yang dihasilkan model, dilakukan analisis menggunakan metode *SHapley Additive exPlanations* (SHAP) sehingga kontribusi setiap fitur terhadap hasil prediksi dapat dijelaskan secara lebih transparan dan mudah diinterpretasikan.



Gambar 4.1. Confusion Matrix

Berdasarkan Gambar 4.1, model *Random Forest* mampu mengklasifikasikan sebagian besar data dengan benar. Dari total 95.972 data aktual kelas negatif, sebanyak 78.212 data (81,5%) berhasil diprediksi sebagai kelas negatif (*True Negative*), sedangkan 17.760 data (18,5%) salah diprediksi sebagai kelas positif (*False Positive*).

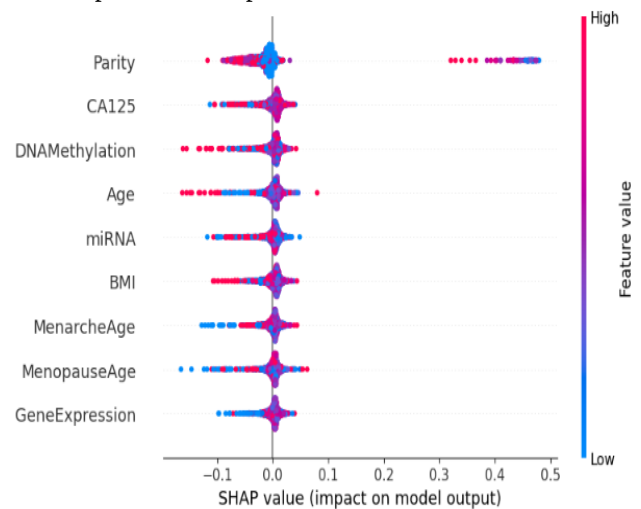
Pada kelas positif, dari total 95.972 data aktual, sebanyak 92.116 data (96,0%) berhasil diprediksi sebagai kelas positif (*True Positive*), sedangkan 3.856 data (4,0%) salah diprediksi sebagai kelas negatif (*False Negative*). Hasil ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengenali data positif, yang ditunjukkan oleh nilai *recall* sebesar 0,94 pada kelas positif.

Secara keseluruhan, hasil *Confusion Matrix* menunjukkan bahwa model *Random Forest* memiliki performa klasifikasi yang baik setelah dilakukan penyeimbangan data menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE). Meskipun masih terdapat kesalahan klasifikasi pada sebagian data, model lebih mampu mengidentifikasi kelas positif dibandingkan kelas negatif. Hal ini penting dalam penelitian prediksi kanker ovarium karena kemampuan mendeteksi kasus positif secara tepat dapat membantu mengurangi risiko terlewatnya pasien yang berpotensi mengalami kanker ovarium.

4.2 Analisis Explainable Artificial Intelligence (XAI) Menggunakan SHAP

Setelah model *Random Forest* dievaluasi, dilakukan analisis menggunakan metode *SHapley Additive exPlanations* (SHAP) untuk meningkatkan transparansi model. SHAP digunakan untuk menjelaskan kontribusi setiap fitur terhadap hasil prediksi sehingga proses pengambilan keputusan model tidak hanya menghasilkan nilai prediksi, tetapi juga memberikan informasi mengenai fitur-fitur yang paling berpengaruh. Analisis SHAP pada penelitian ini terdiri atas *Global Explanation*. *Global*

Explanation digunakan untuk mengetahui pengaruh setiap fitur terhadap keseluruhan model.



Gambar 4.2. Global Explanation

Berdasarkan Gambar 4.2, hasil *SHAP Summary Plot* menunjukkan bahwa Parity merupakan fitur yang paling berpengaruh terhadap prediksi kanker ovarium, diikuti oleh CA125, DNAMethylation, Age, miRNA, BMI, MenarcheAge, MenopauseAge, dan GeneExpression. Semakin tinggi posisi suatu fitur pada grafik, semakin besar kontribusinya terhadap keputusan model. Warna merah menunjukkan nilai fitur yang tinggi, sedangkan warna biru menunjukkan nilai fitur yang rendah. Hasil *Global Explanation* menunjukkan bahwa model *Random Forest* mampu mengidentifikasi fitur-fitur yang memberikan kontribusi terbesar terhadap prediksi sehingga meningkatkan transparansi dan kemudahan interpretasi model.

4.3 Implementasi Sistem

Sistem Prediksi Risiko Kanker Ovarium

Sistem ini membantu melakukan analisis risiko kanker ovarium berdasarkan faktor reproduksi, faktor klinis, dan biomarker molekuler menggunakan algoritma Random Forest dan SHAP.

Silakan isi seluruh pertanyaan sesuai kondisi Anda.

Data Reproduksi

Berapa usia Anda saat ini?

Berapa kali Anda pernah melahirkan?

Pada usia berapa Anda mengalami menstruasi pertama kali (Menarche)?

Pada usia berapa Anda mengalami menopause?

Data Klinis

Berapa nilai BMI Anda saat ini?

Berapa kadar CA-125 berdasarkan hasil laboratorium (U/mL)?

Kosongkan jika belum pernah diperiksa

Gambar 4.3. Halaman Input Prediksi Risiko Kanker Ovarium

Data Klinis

Berapa nilai BMI Anda saat ini?

Berapa kadar CA-125 berdasarkan hasil laboratorium (U/mL)?

Kosongkan jika belum pernah diperiksa

Biomarker Molekuler

Jika Anda belum memiliki hasil pemeriksaan laboratorium, kolom berikut boleh dikosongkan. Fisk akan mengisi nilai rata-rata (mean) dari data pelatihan.

Nilai Gene Expression

Kosongkan jika belum diperiksa

Nilai DNA Methylation

Kosongkan jika belum diperiksa

Nilai miRNA

Kosongkan jika belum diperiksa

Analisis Risiko Kanker Ovarium

Gambar 4.4. Halaman Input Prediksi Risiko Kanker Ovarium

Berdasarkan Gambar 4.3 dan Gambar 4.4 halaman input digunakan untuk memasukkan data pasien yang terdiri atas Data Reproduksi, Data Klinis, dan Biomarker Molekuler. Data reproduksi meliputi usia, jumlah persalinan (*parity*), usia menarche, dan usia menopause. Data klinis meliputi nilai BMI dan CA-125, sedangkan biomarker molekuler terdiri atas Gene Expression, DNA Methylation, dan miRNA.

Apabila pengguna belum memiliki hasil pemeriksaan biomarker molekuler, kolom tersebut dapat dikosongkan dan sistem akan menggunakan nilai rata-rata (*mean*) dari data pelatihan sebagai nilai pengganti. Setelah seluruh data diisi, pengguna dapat menekan tombol Analisis Risiko Kanker Ovarium untuk menjalankan proses prediksi menggunakan algoritma *Random Forest*.

Hasil Analisis Risiko Kanker Ovarium

Data yang Dimasukkan

Parameter	Nilai
Usia	58.0
Jumlah Persalinan	1.0
BMI	22.9
CA-125	0.003166754211142823
Gene Expression	0.0011688995896612707
DNA Methylation	4.167800786864144e-05
miRNA	0.0005987222147098017
Usia Menarche	11.0
Usia Menopause	50.0

Hasil Prediksi
Risiko Tinggi

Probabilitas Risiko

64.93%

Penjelasan SHAP (Explainable AI)

Berdasarkan analisis SHAP, fitur yang paling meningkatkan risiko adalah Jumlah Persalinan dan DNA Methylation. Sementara itu, fitur yang paling menurunkan risiko adalah Usia dan Usia Menopause. Secara keseluruhan, kombinasi kontribusi seluruh fitur menghasilkan prediksi risiko tinggi pada model Random Forest.

Gambar 4.5. Halaman Hasil Analisis Kanker Ovarium

Fitur	Nilai SHAP
Jumlah Persalinan	0.3240
Usia	-0.0938
Usia Menopause	-0.0626
Usia Menarche	-0.0226
DNA Methylation	0.0088
miRNA	-0.0038
BMI	0.0032
Gene Expression	-0.0021
CA-125	-0.0015

[Kembali ke Halaman Utama](#)

Gambar 4.6. Halaman Hasil Analisis Kanker Ovarium

Berdasarkan Gambar 4.5, halaman hasil prediksi menampilkan keluaran sistem setelah pengguna mengisi seluruh data pasien pada formulir prediksi. Pengujian dilakukan menggunakan data pasien sebagai contoh implementasi model *Random Forest* dalam memprediksi risiko kanker ovarium. Berdasarkan data yang dimasukkan, sistem mengklasifikasikan pasien ke dalam kategori Risiko Tinggi dengan probabilitas sebesar 64,93%.

Selain menampilkan hasil prediksi dan nilai probabilitas, sistem juga menyajikan penjelasan menggunakan metode SHapley Additive exPlanations (SHAP). Penjelasan tersebut menunjukkan fitur-fitur yang memberikan kontribusi terbesar dalam meningkatkan maupun menurunkan prediksi risiko kanker ovarium, sehingga pengguna tidak hanya memperoleh hasil prediksi, tetapi juga dapat memahami faktor-faktor yang memengaruhi keputusan model secara lebih transparan.

5. KESIMPULAN

Berdasarkan hasil penelitian, algoritma *Random Forest* berhasil digunakan untuk memprediksi risiko kanker ovarium berdasarkan dataset medis terbuka dengan performa klasifikasi yang baik. Analisis menggunakan metode *SHapley Additive exPlanations* (SHAP) mampu menjelaskan kontribusi setiap fitur terhadap hasil prediksi sehingga meningkatkan transparansi dan interpretabilitas model. Model yang dibangun juga berhasil diimplementasikan dalam aplikasi berbasis web yang dapat menampilkan hasil prediksi, nilai probabilitas, serta penjelasan kontribusi fitur melalui SHAP. Hasil penelitian ini menunjukkan bahwa kombinasi *Random Forest* dan SHAP berpotensi menjadi alat bantu dalam analisis risiko kanker ovarium yang lebih mudah dipahami oleh pengguna.

6. SARAN

Penelitian selanjutnya diharapkan dapat menggunakan dataset yang lebih besar dan beragam agar kemampuan generalisasi model semakin baik. Selain itu, dapat dilakukan perbandingan dengan algoritma *machine learning* lain, seperti XGBoost, LightGBM, *Support Vector Machine* (SVM), atau

Neural Network, untuk memperoleh performa yang lebih optimal. Pengembangan sistem menggunakan data medis nyata dari rumah sakit atau institusi kesehatan serta validasi lebih lanjut pada lingkungan klinis juga diperlukan agar hasil penelitian lebih representatif dan dapat mendukung penerapan sistem sebagai alat bantu pengambilan keputusan di bidang kesehatan

7. REFERENSI

- An, J., Kim, H., Lee, S., & Park, J. (2023). *A review of electronic health records and their applications in machine learning for healthcare*. *Journal of Medical Systems*, 47(1), 1–12.
- Berek, J. S., Matulonis, U. A., Peen, U., Gaarenstroom, K., Colombo, N., Herzog, T. J., & Ledermann, J. A. (2021). *Cancer of the ovary, fallopian tube, and peritoneum: 2021 update*. *International Journal of Gynecology & Obstetrics*, 155(Suppl. 1), 61–85. <https://doi.org/10.1002/ijgo.13878>
- Efthimiou, O., Seo, M., Chalkou, K., Debray, T. P. A., Egger, M., & Salanti, G. (2024). *Developing clinical prediction models: A step-by-step guide*. *BMJ*, 386, e078276. <https://doi.org/10.1136/bmj-2023-078276>
- Fidayanthie, E. W., Rahman, A., Putri, D. A., & Hidayat, R. (2025). *Analisis perbandingan algoritma Random Forest dan algoritma Naive Bayes untuk memprediksi penyakit paru-paru di Indonesia*. *Saturnus: Jurnal Teknologi dan Sistem Informasi*, 3(3), 123–135. <https://doi.org/10.61132/saturnus.v3i3.956>
- Mersha, M., Alemu, T., Kebede, A., & Bekele, G. (2024). *Explainable artificial intelligence: A survey of needs, techniques, applications, and future direction*. *Neurocomputing*. <https://doi.org/10.48550/arXiv.2409.00265>
- Ponce-Bobadilla, J., García, J., Rodríguez, M., & López, A. (2024). *Explainable artificial intelligence (XAI) methods in machine learning: A review*. *Applied Sciences*, 14(3), Article 1234. <https://doi.org/10.3390/app14031234>
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). *Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249. <https://doi.org/10.3322/caac.21660>