

ANALISIS HUBUNGAN ADOPSI ROBOT TERHADAP PENINGKATAN PRODUKTIVITAS MENGGUNAKAN PENDEKATAN RANDOM FOREST CLASSIFIER

Afra Ray Sinta

Politeknik Negeri Batam, Batam, Indonesia

INFORMASI ARTIKEL	ABSTRAK
Sejarah Artikel: Diterima: Juni 2026 Revisi: Juni 2026 Diterima: Juli 2026 Dipublikasi: Juli 2026 Kata Kunci: 1; Random Forest Classifier 2; Otomasi 3; Adopsi Robot *Penulis Korespondensi: afraraysinta2010@gmail.com	Penelitian ini bertujuan untuk menguji secara empiris apakah jumlah robot yang diadopsi memiliki hubungan yang signifikan terhadap peningkatan produktivitas menggunakan pendekatan algoritma <i>artificial intelligence</i>. Data yang digunakan terdiri atas 27 observasi dari tiga sektor industri, yaitu Manufaktur, Kesehatan, dan Logistik, pada rentang tahun 2015 hingga 2023. Variabel <i>productivity gain</i> bersifat kontinu ditransformasikan menjadi variabel kelas biner berdasarkan nilai median, kemudian diklasifikasikan menggunakan algoritma random forest classifier dengan fitur pendukung berupa variabel <i>industry</i>, <i>cost savings</i>, <i>jobs displaced</i>, dan <i>training_hours</i>. Hasil penelitian menunjukkan bahwa model memiliki akurasi pengujian sebesar 55,6 % dan rata rata akurasi validasi silang hanya 26 %, yang berada di bawah tingkat akurasi tebakan acak. Temuan ini mengindikasikan bahwa, berdasarkan data yang tersedia, tidak ditemukan bukti kuat bahwa peningkatan jumlah adopsi robot berasosiasi secara langsung dengan peningkatan produktivitas.

PENDAHULUAN

Perkembangan teknologi otomasi dan robotika telah mendorong berbagai sektor industri, termasuk manufaktur, kesehatan, dan logistik, untuk mengadopsi robot sebagai bagian dari proses operasional mereka. Adopsi robot umumnya diyakini dapat meningkatkan efisiensi kerja, mengurangi kesalahan manusia, serta meningkatkan produktivitas secara keseluruhan. Namun demikian, hubungan antara jumlah robot yang diadopsi dan besarnya peningkatan produktivitas yang dihasilkan belum tentu bersifat linear atau signifikan, karena banyak faktor lain seperti pelatihan tenaga kerja, efisiensi biaya, dan karakteristik sektor industri yang turut memengaruhi hasil akhir.

Penelitian ini bertujuan untuk menjawab pertanyaan penelitian: apakah semakin banyak robot yang diadopsi oleh suatu organisasi akan meningkatkan produktivitasnya secara signifikan? Untuk menjawab pertanyaan tersebut, digunakan pendekatan kecerdasan buatan berupa algoritma Random Forest Classifier, yang dipilih karena kemampuannya menangani data tabular berukuran kecil,

ketahanannya terhadap overfitting melalui mekanisme bagging, serta kemampuannya menghasilkan skor feature importance yang dapat digunakan untuk mengidentifikasi variabel paling berpengaruh terhadap target klasifikasi.

Kontribusi penelitian ini adalah menyediakan bukti empiris berbasis data mengenai kekuatan hubungan antara adopsi robot dan produktivitas, sekaligus mendemonstrasikan tahapan lengkap analisis data menggunakan metode klasifikasi, mulai dari praproses data, ekstraksi fitur, pelatihan model, hingga evaluasi hasil menggunakan confusion matrix, kurva ROC, histogram, dan visualisasi sebaran data.

LANDASAN TEORI

2.1 Adopsi Robot dan Produktivitas

Adopsi robot merupakan bagian dari transformasi teknologi otomasi yang bertujuan menggantikan atau mendukung tugas-tugas yang sebelumnya dikerjakan oleh manusia dalam proses produksi maupun layanan. Secara teoretis, otomasi dapat meningkatkan produktivitas melalui peningkatan kecepatan proses, konsistensi kualitas, dan pengurangan kesalahan operasional. Namun, beberapa kajian menunjukkan bahwa hubungan antara adopsi robot dan produktivitas tidak selalu bersifat linear, karena dipengaruhi oleh faktor lain seperti kesiapan tenaga kerja, jenis industri, dan skala investasi. Acemoglu dan Restrepo mengemukakan bahwa dampak robot terhadap pasar tenaga kerja dan produktivitas bervariasi tergantung pada sektor dan tingkat substitusi tenaga kerja oleh mesin, sehingga peningkatan jumlah robot tidak serta-merta menjamin peningkatan produktivitas secara proporsional [2].

2.2 Klasifikasi dalam Kecerdasan Buatan

Klasifikasi merupakan salah satu tugas utama dalam pembelajaran mesin (*machine learning*) yang bertujuan memetakan sekumpulan variabel prediktor ke dalam kelas atau kategori tertentu. Dalam konteks penelitian ini, variabel target yang semula bersifat kontinu ditransformasikan menjadi variabel kelas biner agar dapat dianalisis menggunakan pendekatan klasifikasi. Pendekatan ini umum digunakan ketika peneliti ingin mengetahui apakah suatu observasi termasuk ke dalam kategori tinggi atau rendah berdasarkan ambang batas tertentu, seperti nilai median.

2.3 Random Forest Classifier

Random Forest Classifier merupakan metode pembelajaran ensemble yang diperkenalkan oleh Breiman, yang bekerja dengan cara membangun sejumlah besar decision tree secara acak melalui teknik bootstrap aggregating (bagging) dan menggabungkan hasil prediksi masing-masing pohon melalui mekanisme voting mayoritas. Setiap pohon dilatih menggunakan subset data dan subset fitur yang dipilih secara acak, sehingga metode ini relatif tahan terhadap overfitting dibandingkan decision tree tunggal, terutama pada dataset berukuran kecil. Selain menghasilkan prediksi kelas, Random Forest juga mampu menghitung nilai feature importance, yaitu ukuran kontribusi relatif setiap variabel prediktor terhadap akurasi klasifikasi, sehingga dapat digunakan untuk mengidentifikasi variabel yang paling berpengaruh terhadap variabel target.

2.4 Validasi Silang dan Evaluasi Model

Validasi silang (cross-validation) merupakan teknik evaluasi model yang bertujuan mengukur kemampuan generalisasi model terhadap data baru dengan cara membagi data menjadi beberapa lipatan (fold), melatih model pada sebagian lipatan, dan mengujinya pada lipatan yang tersisa secara bergantian. Teknik ini penting digunakan terutama pada dataset berukuran kecil untuk mengurangi bias akibat pembagian data latih dan data uji yang hanya dilakukan sekali. Evaluasi performa model klasifikasi umumnya dilakukan menggunakan confusion matrix serta kurva Receiver Operating Characteristic (ROC) beserta nilai Area Under Curve (AUC). Confusion matrix menggambarkan jumlah prediksi benar dan salah untuk setiap kelas, sedangkan kurva ROC dan AUC digunakan untuk menilai kemampuan model dalam membedakan antar kelas pada berbagai ambang batas probabilitas, di mana nilai AUC mendekati 0,5 mengindikasikan kemampuan diskriminasi model yang setara dengan tebakan acak.

METODE PENELITIAN

2.1 Data Penelitian

Data yang digunakan dalam penelitian ini berjumlah 27 observasi yang mencakup periode tahun 2015 hingga 2023 pada tiga sektor industri, yaitu Manufaktur, Kesehatan, dan Logistik. Variabel yang tersedia dalam dataset meliputi Year, Industry, Robots_Adopted, Productivity_Gain, Cost_Savings, Jobs_Displaced, dan Training_Hours. Variabel Robots_Adopted merupakan variabel bebas utama yang merepresentasikan jumlah robot yang diadopsi, sedangkan Productivity_Gain merupakan variabel target yang merepresentasikan persentase peningkatan produktivitas.

2.2 Penentuan Metode Kecerdasan Buatan

Metode yang dipilih dalam penelitian ini adalah Random Forest Classifier, yaitu metode pembelajaran ensemble yang membangun banyak decision tree secara acak (bootstrap aggregating) dan menggabungkan hasil prediksinya melalui mekanisme voting. Metode ini dipilih dengan pertimbangan sebagai berikut: (1) sesuai untuk data tabular dengan kombinasi fitur numerik dan kategorik; (2) relatif tahan terhadap overfitting meskipun jumlah data kecil; dan (3) mampu menghasilkan nilai feature importance yang berguna untuk mengidentifikasi variabel yang paling berkontribusi terhadap klasifikasi produktivitas.

2.3 Praproses Data

- Pemeriksaan missing value: tidak ditemukan nilai yang hilang pada seluruh variabel.
- Label encoding pada variabel kategorik Industry menjadi representasi numerik.
- Transformasi variabel target Productivity_Gain (kontinu) menjadi variabel kelas biner (0 = Low, 1 = High) berdasarkan nilai median sebesar 15,71.
- Standardisasi (z-score) pada seluruh fitur numerik menggunakan StandardScaler.

2.4 Ekstraksi Fitur

Fitur yang digunakan dalam proses klasifikasi terdiri atas Robots_Adopted sebagai variabel utama penelitian, serta Industry (terencode), Cost_Savings, Jobs_Displaced, dan Training_Hours sebagai variabel pendukung yang berpotensi turut memengaruhi produktivitas.

2.5 Klasifikasi

Data dibagi menjadi data latih (70%) dan data uji (30%) menggunakan teknik stratified split agar proporsi kelas tetap seimbang. Model Random Forest dilatih dengan 200 pohon keputusan ($n_estimators = 200$) dan kedalaman maksimum 4 ($max_depth = 4$) untuk mencegah overfitting pada data yang berukuran kecil. Untuk memperoleh estimasi performa yang lebih stabil, dilakukan pula validasi silang (cross-validation) dengan skema stratified 5-fold

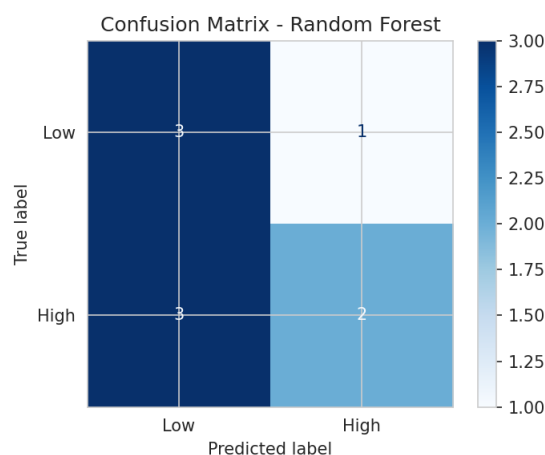
2.6 Evaluasi Model

Evaluasi hasil klasifikasi dilakukan menggunakan confusion matrix, kurva ROC beserta nilai Area Under Curve (AUC), histogram distribusi Productivity_Gain berdasarkan kelas, diagram sebar (scatter plot) antara Robots_Adopted dan Productivity_Gain, serta analisis feature importance untuk mengidentifikasi variabel yang paling berpengaruh terhadap hasil klasifikasi.

HASIL DAN PEMBAHASAN

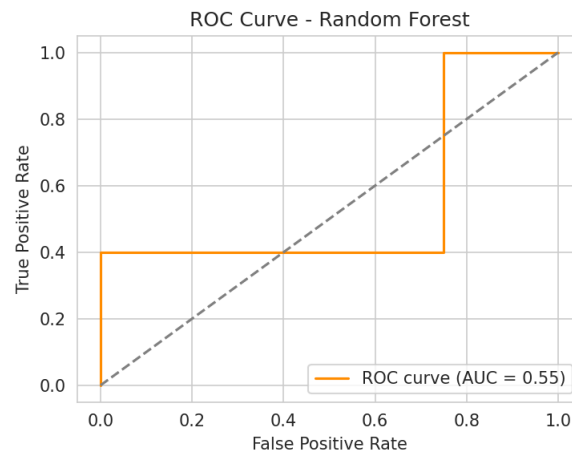
Performa Model Klasifikasi

Model Random Forest yang dilatih menghasilkan akurasi sebesar 55,6% pada data uji. Sementara itu, hasil validasi silang 5-fold menunjukkan rata-rata akurasi yang jauh lebih rendah, yaitu 26%, dengan standar deviasi 0,17. Nilai ini berada di bawah tingkat akurasi tebakan acak untuk klasifikasi dua kelas (50%), yang mengindikasikan bahwa model tidak mampu menemukan pola prediktif yang konsisten dari fitur-fitur yang tersedia terhadap kelas produktivitas.



Gambar 1. Confusion Matrix hasil klasifikasi Random Forest pada data uji

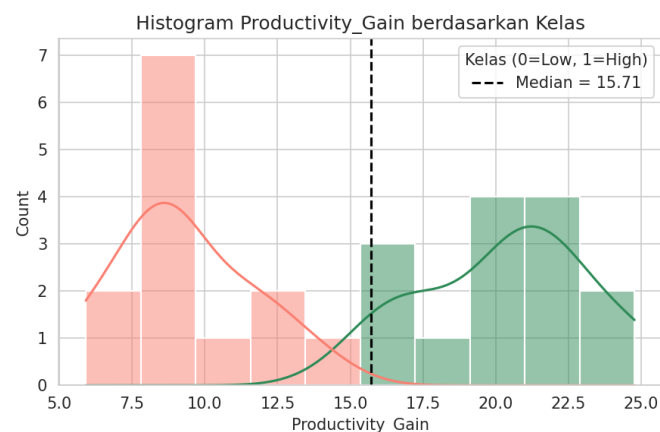
Confusion matrix pada Gambar 1 menunjukkan bahwa model masih melakukan cukup banyak kesalahan klasifikasi antara kelas Low dan High, yang memperkuat indikasi lemahnya pola prediktif dalam data.



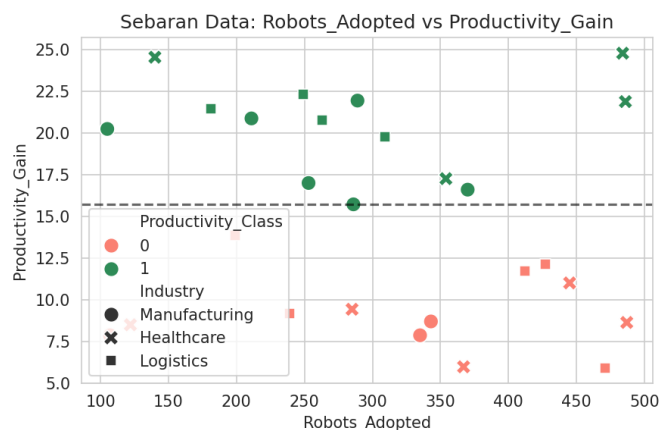
Gambar 2. Kurva ROC dan nilai AUC model Random Forest

Kurva ROC pada Gambar 2 menggambarkan kemampuan model dalam membedakan kelas High dan Low pada berbagai ambang batas probabilitas. Kurva yang mendekati garis diagonal (AUC mendekati 0,5) menunjukkan bahwa kemampuan diskriminatif model tidak jauh berbeda dari klasifikasi acak.

Distribusi dan Sebaran Data



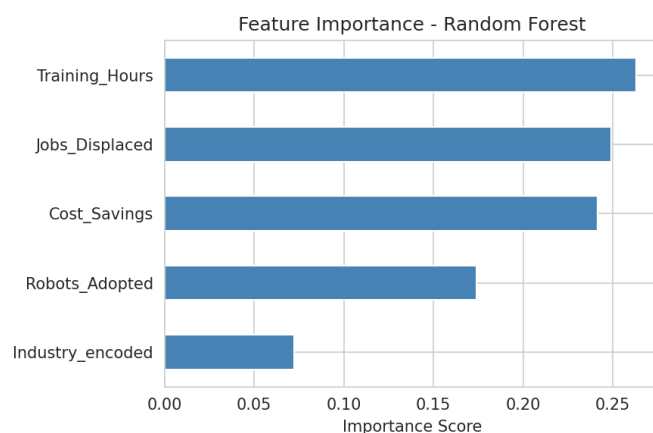
Histogram pada Gambar 3 memperlihatkan bahwa distribusi nilai Productivity_Gain pada kedua kelas relatif tumpang tindih (overlap) di sekitar nilai median, tanpa pemisahan yang jelas, yang menunjukkan bahwa variabel-variabel prediktor yang digunakan belum mampu membedakan kedua kelas secara tegas.



Gambar 4. Sebaran data Robots_Adopted terhadap Productivity_Gain berdasarkan kelas dan sektor industri

Diagram sebar pada Gambar 4 menunjukkan tidak adanya tren atau pola yang jelas antara jumlah robot yang diadopsi dan besarnya peningkatan produktivitas. Titik data dengan jumlah robot yang tinggi maupun rendah tersebar merata pada kedua sisi garis median, baik untuk kelas Low maupun High, pada ketiga sektor industri yang diteliti.

Analisis Feature Importance



Gambar 5. Skor feature importance dari model Random Forest

Berdasarkan Gambar 5, urutan tingkat kepentingan fitur dari yang tertinggi adalah Training_Hours (0,263), Jobs_Displaced (0,249), Cost_Savings (0,242), Robots_Adopted (0,174), dan Industry (0,072). Hasil ini menunjukkan bahwa Robots_Adopted, yang merupakan variabel utama dalam pertanyaan penelitian, justru bukan merupakan fitur dengan kontribusi tertinggi terhadap klasifikasi kelas produktivitas.

Fitur	Skor Feature Importance
Training_Hours	0,263

Jobs_Displaced	0,249
Cost_Savings	0,242
Robots_Adopted	0,174
Industry	0,072

Tabel 1. Ringkasan skor feature importance model Random Forest

Analisis Korelasi

Analisis korelasi Pearson antara Robots_Adopted dan Productivity_Gain menghasilkan nilai $r = -0,173$, yang menunjukkan hubungan linear yang sangat lemah dan bahkan berarah negatif. Nilai korelasi ini memperkuat hasil dari model klasifikasi maupun analisis feature importance, bahwa penambahan jumlah robot yang diadopsi tidak serta-merta diikuti oleh peningkatan produktivitas dalam dataset yang diteliti.

Pembahasan

Secara keseluruhan, seluruh indikator analisis, mulai dari akurasi klasifikasi, confusion matrix, kurva ROC, histogram distribusi, diagram sebar, feature importance, hingga korelasi Pearson, menunjukkan arah kesimpulan yang konsisten, yaitu tidak ditemukannya hubungan yang kuat antara jumlah robot yang diadopsi dan besarnya peningkatan produktivitas pada dataset yang digunakan. Temuan ini berbeda dengan asumsi umum bahwa otomasi selalu berbanding lurus dengan produktivitas. Kemungkinan penjelasan atas temuan ini antara lain: (1) peningkatan produktivitas kemungkinan lebih dipengaruhi oleh faktor lain yang tidak tercakup dalam dataset, seperti kualitas pelatihan, jenis robot, atau tingkat integrasi proses; dan (2) kemungkinan data yang digunakan bersifat simulatif sehingga tidak merepresentasikan hubungan sebab-akibat yang sesungguhnya terjadi di lapangan.

KESIMPULAN

Berdasarkan hasil analisis menggunakan Random Forest Classifier terhadap data adopsi robot dan produktivitas pada 27 observasi dari tiga sektor industri, dapat disimpulkan bahwa tidak ditemukan bukti yang kuat mengenai hubungan antara jumlah robot yang diadopsi (Robots_Adopted) dan peningkatan produktivitas (Productivity_Gain). Hal ini ditunjukkan oleh rendahnya akurasi model (55,6% pada data uji dan 26% pada validasi silang), lemahnya korelasi Pearson ($r = -0,173$), serta rendahnya kontribusi Robots_Adopted pada analisis feature importance dibandingkan variabel lain seperti Training_Hours, Jobs_Displaced, dan Cost_Savings.

Penelitian selanjutnya disarankan untuk menggunakan jumlah data yang lebih besar serta menambahkan variabel-variabel lain yang secara teoretis lebih relevan, seperti jenis dan kompleksitas robot, tingkat integrasi otomasi dalam proses produksi, kompetensi tenaga kerja, maupun karakteristik organisasi, guna memperoleh gambaran yang lebih komprehensif mengenai hubungan antara otomasi dan produktivitas.

Daftar Pustaka

- [1] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- [2] Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, 128(6), 2188–2244.
- [3] Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- [4] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874.