

PENERAPAN METODE HYBRID RANDOM FOREST DAN GENETIC ALGORITHM UNTUK OPTIMASI PENJADWALAN PRODUKSI

Widya Monika Sari^{1*}, Nurmalitasari², Bangun Prajadi Cipto Utomo³

^{1,2,3} Universitas Duta Bangsa Surakarta, Indonesia

^{1*}220103253@mhs.udb.ac.id, ²nurmalitasari@udb.ac.id, ³bangun_prajadi@udb.ac.id

DOI: <https://doi.org/10.46880/mtk.v12i2.5789>

ABSTRACT

The garment industry faces complex production scheduling challenges due to high product variability and inaccurate process time estimation. Inefficient production scheduling often leads to production delays and reduced operational performance. Conventional scheduling methods such as First Come First Serve (FCFS) and Earliest Due Date (EDD) have been shown to be less effective in dynamic production environments. This study aims to optimize flow shop production scheduling in a children's garment manufacturing environment using a hybrid Random Forest–Genetic Algorithm approach. Random Forest is employed to predict the processing time of each job based on a simulated dataset regenerated from the company's historical production data collected in 2025 while preserving the statistical characteristics and relationships among variables. Subsequently, the Genetic Algorithm is used to optimize job sequencing by simultaneously minimizing makespan and weighted tardiness. The study follows the CRISP-DM methodology up to the model evaluation stage. The results show that the Random Forest model achieved satisfactory prediction performance for the cutting, sewing, and finishing stages, with R^2 values of 0.817, 0.981, and 0.867, respectively. Using a population size of 30, 100 generations, and 10 independent runs, the Genetic Algorithm achieved an improvement of 79.94% compared to FCFS and 39.71% compared to EDD. These findings demonstrate that the proposed hybrid Random Forest–Genetic Algorithm approach can generate a more adaptive, efficient, and data-driven production schedule than conventional scheduling methods.

Keywords: Genetic Algorithm, Flow Shop Scheduling, Scheduling Optimization, Random Forest, CRISP-DM

I. PENDAHULUAN

Perusahaan Garmen X merupakan perusahaan yang bergerak pada sektor manufaktur produksi pakaian anak seperti kaos, dress, celana, topi, dll. Dalam proses produksi tiap produk harus melewati beberapa tahapan produksi yang saling terhubung mulai dari *cutting*, *sewing*, *finishing* hingga *packing* [1]. Kompleksitas proses produksi akan meningkat seiring bertambahnya variasi produk dan jumlah permintaan produksi sehingga perusahaan harus mampu menghasilkan produk secara cepat dan tepat waktu. Salah satu aspek yang mempengaruhi ketepatan proses produksi salah satunya penjadwalan produksi. Penjadwalan produksi merupakan proses pengalokasian sumber daya seperti waktu, mesin, dan tenaga kerja bertujuan untuk efisiensi dan meminimalkan keterlambatan proses produksi [2].

Permasalahan penjadwalan produksi merupakan permasalahan yang kompleks karena melibatkan banyak variabel kombinatorial yang berkaitan dengan penentuan urutan pekerjaan (*job sequencing*) serta alokasi sumber daya yang terbatas [3]. Penjadwalan modern tidak hanya bergantung pada aturan statis, tetapi diperlukan pemahaman dalam pola data historis serta kemampuan prediksi parameter penting seperti waktu proses dan keterlambatan produksi [4]. Sistem produksi dengan pola *flow shop scheduling* memiliki karakteristik saling berkaitan dimana setiap *job* di proses berurutan dari proses awal sampai akhir tanpa adanya variasi jalur produksi. Hal ini menjadikan parameter waktu produksi sangat penting dalam optimasi penjadwalan produksi.

Namun dalam melakukan penjadwalan, perusahaan X masih menggunakan metode konvensional seperti *first come first serve* (FCFS) dan

earliest due date (EDD). Metode tersebut relatif sederhana dan mudah diterapkan, tetapi sering kali belum mampu menghasilkan solusi optimal ketika dihadapkan pada kondisi produksi yang dinamis dengan jumlah *job* dan variasi produk tinggi [5]. Penjadwalan masih sangat bergantung pada pengalaman kepala produksi atau mandor sehingga keputusan yang dihasilkan kurang Objektif dan kurang adaptif terhadap perubahan kondisi produksi. Akibatnya sering terjadi keterlambatan penyelesaian pekerjaan, penumpukan *job* pada stasiun kerja tertentu, serta waktu proses produksi yang kurang efisien. Selain itu data historis produksi belum dimanfaatkan secara optimal sebagai dasar pengambilan keputusan dalam proses [6]. Akibatnya, penjadwalan berjalan di atas asumsi yang tidak selalu mencerminkan realitas lantai produksi.

Dalam praktik industri garmen, sebelum proses produksi dilakukan, perusahaan umumnya menerapkan *time study* atau pengukuran waktu kerja untuk memperoleh estimasi durasi proses pada setiap tahapan produksi. Pengukuran tersebut dilakukan berdasarkan kecepatan operator, jenis pekerjaan, tingkat kesulitan produk, serta kapasitas mesin pada tiap stasiun kerja, khususnya pada proses *sewing* yang memiliki banyak aktivitas jahit dengan tingkat kompleksitas berbeda. Estimasi waktu hasil *time study* digunakan sebagai acuan awal dalam perencanaan produksi, namun pada kondisi nyata sering terjadi perbedaan antara estimasi dan realisasi waktu produksi akibat faktor operasional yang dinamis. Oleh karena itu, diperlukan pendekatan yang mampu mempelajari pola data historis dan melakukan koreksi terhadap estimasi waktu produksi agar penjadwalan menjadi lebih akurat dan adaptif.



Berbagai penelitian sebelumnya yang telah mengkaji optimasi penjadwalan produksi menggunakan pendekatan *metaheuristic* maupun *machine learning*[7]. Seperti penelitian optimasi penjadwalan yang telah dilakukan oleh Anala Agusnur (2025) menggunakan algoritma genetika untuk optimasi penjadwalan produksi pada sistem manufaktur untuk menurunkan *makespan*[8]. Kemudian pada penelitian terkait penggunaan metode hybrid RF-GA seperti yang dilakukan oleh Hashemi dkk (2022) menggunakan model *hybrid random forest-genetic algorithm* untuk memprediksi estimasi radiasi matahari harian menggunakan data meteorologi [9]. Serta Norouzi dkk (2023) menggunakan metode *hybrid GA-RF* untuk sistem deteksi intrusi pada IoMT [10]. Pada konteks penjadwalan produksi Lili Dai dkk (2022) mengembangkan *improvement hybrid genetic algorithm* (IHGA) yang mampu mengatasi kelemahan GA konvensional [7]. Pada konteks *flow shop scheduling*, Aslihan Cubukcuoglu dkk (2025) menggunakan metode PSA-Genetic Algorithm untuk meminimalkan *makespan* pada masalah penjadwalan *flow shop*[11]. Penelitian serupa dilakukan oleh Yu-Yan Zhang dkk (2025) menggunakan RF-GA untuk mengembangkan metode tunggal GA dalam menangani permasalahan kompleks pada penjadwalan *flow shop* untuk mencari solusi terbaik yang efektif [12].

Tabel 1 Penelitian terdahulu

Peneliti	Metode dan Permasalahan	Hasil
Anala Agusnur E [8]	Permasalahan penjadwalan pada industri manufaktur yang kompleks sulit dioptimalkan dengan menggunakan metode konvensional seperti FCFS dan SPT, sehingga diperlukan pendekatan optimasi yang lebih adaptif dengan Algoritma genetika	Algoritma genetika dengan parameter populasi 50, peluang crossover 0,8 dan mutasi 0.5 mampu menurunkan waktu penyelesaian total(<i>makespan</i>) dengan rata-rata 17, 31% dengan waktu komputasi yang efisien
Hashemi dkk [9]	Penelitian ini menggunakan model hybrid Random forest-genetic algorithm untuk prediksi radiasi harian energi surya sebagai energi terbarukan dan merupakan energi utama di bumi.	Model GA-RF-V pada Nir stasiun menunjukkan kinerja terbaik dengan nilai RMSE 0.346 MJ/m ² /hari dan nilai KGE sebesar 0.687
Norouzi dkk[10]	Sistem deteksi intrusi pada IoMT masih perlu ditingkatkan karena masalah keamanan sistem yang rentan terhadap serangan malware. Metode genetic algorithm digunakan untuk mengoptimalkan proses klasifikasi Random forest untuk meningkatkan kemampuan deteksi intrusi sistem.	Evaluasi model hybrid dengan metrik accuracy model mencapai 99.99%, precision 100% dan recall 100%.
Lili Dai dkk[7]	Kompleksitas penjadwalan proses produksi pada industri manufaktur yang tidak dapat dioptimalkan secara efektif oleh algoritma genetika tunggal. Metode gabungan	Hasil dari IHGA memberikan optimasi yang lebih baik sebesar 4.2% - 6.9% dibandingkan metode GA tunggal. Dan rata-rata <i>simulated</i>

	algoritma genetika dan <i>simulated annealing</i> diusulkan untuk meningkatkan pencarian solusi optimal	<i>annealing</i> memberikan peningkatan 5.7% - 8.4% pada IHGA
Aslihan Cubukcuoglu dkk [11]	Penjadwalan produksi pada pola ordered flowshop scheduling yang masih kurang efektif untuk minimasi <i>makespan</i> . Optimasi yang diusulkan menggunakan genetic algorithm yang dimodifikasi dengan <i>pyramid structure adapted genetic algorithm</i>	Hasil analisis statistik menunjukkan bahwa PSA-GA menghasilkan nilai <i>makespan</i> yang lebih kecil dibanding metode NEH, <i>Pair Insert</i> , ILS, dan PSA-TS
Yu-yan Zhang [12]	Penelitian ini mengembangkan metode simple genetic algorithm random forest untuk menyelesaikan masalah order acceptance scheduling dengan pasat sequence dependent setup time. Integrasi random forest sebagai surrogate model untuk memprediksi kualitas solusi sebelum evaluasi dilakukan.	Hasil eksperimen pada 160 benchmark menunjukan SGARF menghasilkan profit tertinggi dengan peningkatan 5.6% dibanding SABPO,SGA, dan SGALS.

Algoritma genetik banyak digunakan untuk mencari solusi urutan *job* optimal karena mampu menyelesaikan permasalahan optimasi yang kompleks[13]. Namun Sebagian besar penelitian masih menggunakan data waktu proses statis atau estimasi manual sehingga kurang mampu merepresentasikan kondisi produksi yang sebenarnya. Metode *machine learning* seperti *random forest* telah banyak dimanfaatkan karena memiliki kemampuan yang baik dalam menangani data non-linear dan kompleks[12]. *Random forest* mampu menghasilkan prediksi yang lebih *smooth* dan menyediakan estimasi keuntungan marjinal dengan biaya komputasi yang rendah dibandingkan metode lain yang cenderung mengalami kesulitan dalam menemukan solusi terbaik dan sering terjebak pada area solusi kurang optimal pada kasus berukuran besar [12]. Penelitian ini yang mengintegrasikan model prediksi waktu berbasis ML dengan optimasi penjadwalan produksi pada lingkungan *flow shop* di industri garmen masih relatif terbatas[14]. Gap penelitian inilah yang menjadi dasar penting dilakukannya penelitian ini.

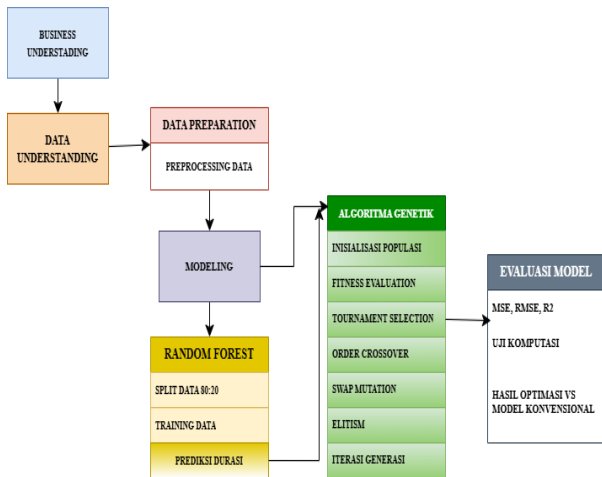
Untuk menjawab permasalahan tersebut, penelitian ini mengusulkan pendekatan *hybrid* yang menggabungkan metode *random forest* sebagai model prediktif dan *genetic algorithm* sebagai metode optimasi. Metode *random forest* difokuskan dalam mengolah dan mempelajari data historis produksi dan koreksi terhadap estimasi waktu hasil *time study* untuk menghasilkan prediksi waktu proses yang lebih akurat dan adaptif terhadap kondisi produksi[15][16]. Hasil prediksi kemudian digunakan sebagai masukan bagi algoritma genetika untuk mencari urutan *job* optimal yang meminimalkan *tardiness*.

Kontribusi utama penelitian ini adalah integrasi model prediktif dan optimasi dalam satu kerangka penjadwalan produksi berbasis data pada industri garmen, sebuah

pendekatan yang tidak hanya mencari urutan *job* terbaik, tetapi juga memastikan bahwa basis waktu proses yang digunakan mencerminkan kondisi produksi nyata. Dengan penelitian ini, diharapkan sistem penjadwalan

II. METODE

Pada Gambar 1 Diagram alur Metode penelitian yang digunakan dalam penelitian ini menggunakan metode *cross industry standart process for data mining* (CRISP-DM). Metode ini dipilih karena menyediakan tahapan sistematis dalam proses pengolahan data, pembangunan model prediksi, hingga implementasi hasil analisis untuk mendukung pengambilan keputusan.



Gambar 1 Skema metode penelitian

Tahapan metode ini meliputi *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*. Penelitian ini tidak mencakup tahapan *deployment*, hal ini karena tujuan utama penelitian adalah mengembangkan serta mengevaluasi kinerja model *hybrid random forest – genetic algorithm* dalam memprediksi durasi produksi dan mengoptimalkan penjadwalan produksi. Pada penelitian ini, *Random Forest* digunakan untuk memprediksi durasi waktu proses produksi tiap *job*, sedangkan algoritma genetik digunakan untuk mengoptimalkan urutan pengerjaan *job* sehingga diperoleh jadwal produksi yang efisien. Fokus penelitian diarahkan pada pengukuran akurasi model prediksi dan efektivitas algoritma optimasi melalui parameter evaluasi. Oleh karena itu, penelitian dibatasi sampai tahap evaluasi untuk memastikan bahwa model yang diusulkan telah memiliki performa yang baik sebelum diimplementasikan dalam sistem secara utuh.

1. Business Understanding

Tahap *Business Understanding* dilakukan untuk mengidentifikasi permasalahan yang terjadi pada proses penjadwalan produksi di Perusahaan Garment X. Berdasarkan hasil observasi dan wawancara, perusahaan masih menggunakan metode penjadwalan konvensional seperti FCFS (*First Come First Serve*) dan EDD (*Earliest Due Date*) yang belum mampu menghasilkan jadwal produksi secara optimal ketika jumlah order meningkat dan variasi produk semakin banyak. Kondisi tersebut menyebabkan terjadinya keterlambatan

produksi, penumpukan pekerjaan pada stasiun kerja tertentu, dan kurang optimalnya pemanfaatan sumber daya produksi. Oleh karena itu penelitian ini bertujuan membangun sistem penjadwalan produksi berbasis *machine learning* dan optimasi yang mampu memprediksi waktu proses produksi serta menghasilkan urutan *job* yang optimal untuk meminimalkan tardiness dan makespan.

2. Data Understanding

Data Understanding dilakukan untuk memahami karakteristik data yang digunakan dalam penelitian. Data yang digunakan merupakan data historis produksi tahun 2025 yang diperoleh dari Perusahaan Garment X. Data awal masih berupa catatan produksi manual yang berisi informasi aktivitas produksi pada berbagai tahapan pekerjaan. Selanjutnya data didigitalisasi ke dalam format Microsoft Excel agar dapat diolah secara komputasional. Atribut yang digunakan dalam penelitian meliputi jumlah produksi, jumlah operator, jumlah mesin, tingkat kesulitan pekerjaan, Pengalaman, jenis bahan dan durasi proses produksi. Setelah proses digitalisasi dilakukan analisis awal terhadap struktur data, distribusi data, serta hubungan antar variabel yang mempengaruhi waktu produksi.

3. Data Preparation

Tahap *Data Preparation* bertujuan mempersiapkan data sebelum digunakan pada proses pemodelan. Tahapan pertama adalah melakukan digitalisasi data historis ke dalam format Excel. Selanjutnya dilakukan pemisahan data berdasarkan tahapan produksi yaitu *Cutting*, *Sewing*, dan *Finishing* sehingga setiap proses memiliki dataset yang terpisah. Total dataset yang digunakan untuk penelitian ini mencakup 1000 data pada setiap tahapan atau 3000 dataset. Setelah itu dilakukan proses *preprocessing* yang meliputi pembersihan data, penanganan *missing value*, transformasi tipe data, dan *encoding* data kategorikal.

Karena jumlah data historis yang tersedia masih terbatas, Penelitian ini menggunakan pendekatan regenerasi dataset simulatif berbasis aturan produksi (*mathematical simulation-based data regeneration*) sebagai solusi untuk menghasilkan data latih yang representatif. Pendekatan simulatif ini tidak bersifat acak sepenuhnya, melainkan dibangun dengan mencakup karakteristik setiap tahapan produksi, faktor yang mempengaruhi durasi pengerjaan, serta variabilitas alami yang terjadi di lapangan. Durasi setiap proses produksi dibangun sebagai fungsi dari beberapa variabel utama yang berpengaruh terhadap waktu penyelesaian pekerjaan seperti jumlah produksi, operator, mesin, dan noise produksi. Data dibangun dengan formula deterministik yang ditambahkan *noise gaussian* untuk merepresentasikan variabilitas alamiah. Secara umum, durasi simulatif produksi untuk setiap *job* ke-*i* dituliskan sebagai berikut :

$$D_i = f(Q_i, F_{model,i}, F_{material,i}, O_i, M_i, E_i, K_i, C_i, G_i, P_i) + \varepsilon_i \quad (1)$$

D_i merupakan durasi simulatif yang diperoleh dari faktor Q_i jumlah produk, $F_{model,i}$ faktor model, $F_{material,i}$ bahan baku, O_i jumlah Operator, M_i jumlah Mesin, E_i tingkat pengalaman operator, K_i tingkat

kesulitan, C_i kondisi mesin, G_i faktor kelompok produksi, P_i Prioritas pesanan dijumlah dengan komponen acak yang merepresentasikan variasi proses produksi.

Untuk setiap job ke- i , waktu dasar produksi dihitung menggunakan hasil *time study* :

$$B_i = \frac{Q_i}{400} \times \theta_s \quad (2)$$

Dimana :

B_i = waktu dasar produksi job ke- i

Q_i = jumlah produk

θ_s = konstanta waktu standar proses

Waktu dasar produksi diperoleh dengan membagi jumlah produk dengan 400 yang merupakan batch standar ukuran lot produksi dan dikali dengan konstanta *time study* tiap tahapan yang berbeda yang berasal dari asumsi *time study* produksi perusahaan.

Noise gaussian digunakan untuk meniru ketidakpastian proses produksi dengan distribusi normal rata-rata 0 dan simpangan baku sebesar 2% dari durasi.

$$D_i = D_i + N(0, 0.02D_i) \quad (3)$$

Standar deviasi 2% dipilih agar terdapat noise realistis tetapi tidak merusak pola utama data. Setiap tahapan produksi memiliki karakteristik distribusi durasi yang berbeda mencerminkan kompleksitas dan jenis pekerjaan yang berbeda pula. Dataset yang dihasilkan kemudian divalidasi untuk memastikan bahwa pola hubungan antar variabel sesuai dengan ekspektasi domain. Validasi mencakup analisis statistik deskriptif, uji korelasi, serta evaluasi distribusi data.

4. Modeling

Tahap *Modeling* terdiri dari dua proses utama untuk membuat model *hybrid* yaitu pembangunan model *Random Forest* untuk prediksi waktu produksi dan optimasi penjadwalan menggunakan *Genetic Algorithm*.

Pemodelan Prediksi *Random Forest*

Tahapan ini dilakukan dengan menggunakan algoritma *random forest* untuk menghasilkan pohon keputusan untuk memberikan hasil prediksi waktu proses setiap *job* dengan melatih model menggunakan data historis produksi. Model RF dibangun secara terpisah untuk setiap tahapan produksi (*Cutting*, *Sewing*, *Finishing*), mengingat setiap tahapan memiliki karakteristik dan pola yang berbeda. Model RF untuk setiap tahapan produksi dikonfigurasi dengan hyperparameter yang dioptimalkan melalui proses validasi silang (*cross-validation*). Proses pelatihan menggunakan skema *train-test split* dengan rasio 80:20 untuk memastikan evaluasi yang objektif. Performa model dievaluasi menggunakan tiga metrik utama yaitu *Mean Absolute Error* (MAE) yang mengukur rata-rata absolut kesalahan prediksi, *Root Mean Squared Error* (RMSE) yang memberikan penalti lebih besar terhadap kesalahan prediksi yang besar, dan koefisien determinasi (R^2) yang mengindikasikan proporsi variansi yang dapat dijelaskan oleh model. *Output* dari tahapan ini adalah matriks waktu proses.

Optimasi Penjadwalan Algoritma Genetika

Penjadwalan produksi pada penelitian ini dimodelkan menggunakan pendekatan *flow shop scheduling*, di

mana setiap *job* diproses secara berurutan melalui tahapan *cutting*, *sewing*, *finishing*. Untuk memperoleh urutan pengerjaan yang optimal, digunakan *Genetic Algorithm* (GA) sebagai metode optimasi dengan tujuan meminimalkan keterlambatan produksi (*tardiness*). Setiap kromosom pada GA merepresentasikan satu kemungkinan urutan *job order* yang dikodekan dalam bentuk permutasi indeks pekerjaan. Populasi awal dibangkitkan secara acak sebanyak 30 kromosom (*population size* = 30) untuk menghasilkan variasi solusi awal yang cukup luas dalam proses pencarian solusi optimal.

Evaluasi setiap kromosom dilakukan menggunakan fungsi *fitness* yang dihitung berdasarkan kombinasi nilai *makespan* dan jumlah pekerjaan terlambat (*late job*). Nilai *tardiness* setiap *job* dihitung menggunakan nilai *completion time* yang dihitung secara kumulatif berdasarkan prediksi durasi produksi yang dihasilkan oleh model *Random Forest* pada setiap tahapan proses produksi. Semakin kecil nilai *fitness* maka solusi penjadwalan dianggap semakin optimal karena mampu meminimalkan keterlambatan dan waktu penyelesaian produksi.

$$Fitness(X) = \alpha \cdot \frac{C_{max}}{C_{max}} + \beta \cdot \frac{\sum w_i T_i}{T_{max}} \quad (4)$$

Fungsi *fitness* menggunakan nilai *makespan* dan *tardiness* yang dinormalisasi agar berada pada skala yang seimbang. Penggunaan bobot pada fungsi *fitness* mengacu pada pendekatan *weighted sum multi-objective optimization* yang banyak digunakan pada permasalahan penjadwalan manufaktur[17].

Proses seleksi menggunakan metode *tournament selection* dengan ukuran turnamen sebanyak 3 kromosom. Metode ini dipilih karena mampu menjaga keseimbangan antara tekanan seleksi dan keberagaman populasi. Tahap *crossover* menggunakan metode *Order Crossover* (OX) yang dirancang khusus untuk representasi permutasi sehingga *offspring* yang dihasilkan tetap valid tanpa adanya duplikasi maupun kehilangan *job order*. Selanjutnya proses mutasi dilakukan menggunakan *swap mutation*, yaitu menukar dua posisi gen secara acak dalam kromosom untuk menjaga keberagaman solusi dan meningkatkan eksplorasi ruang pencarian.

Proses evolusi dijalankan selama 100 generasi (*generation* = 100) pada setiap eksperimen optimasi. Pada setiap generasi, kromosom dengan nilai *fitness* terbaik dipertahankan sebagai kandidat solusi penjadwalan optimal hingga diperoleh urutan produksi yang mampu meminimalkan keterlambatan dan meningkatkan efisiensi sistem produksi.

5. Evaluation

Tahap *Evaluation* dilakukan untuk mengukur performa model *Random Forest* dan *Genetic Algorithm*. Evaluasi *Random Forest* menggunakan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan koefisien determinasi (R^2). Sementara itu, performa *Genetic Algorithm* dievaluasi menggunakan parameter *Total Tardiness*, dan *Late Job*. Untuk menguji konsistensi performa *Genetic Algorithm*, dilakukan

eksperimen komputasi sebanyak 10 kali *running* independen. Setiap *running* menggunakan parameter populasi sebanyak 30 kromosom dan jumlah generasi sebanyak 100. Hasil dari seluruh *running* kemudian dirata-ratakan untuk memperoleh performa optimasi yang lebih representatif. Sebagai pembandingan, metode FCFS dan EDD juga diuji menggunakan data yang sama sehingga dapat diketahui tingkat *improvement* yang diperoleh dari metode yang diusulkan. Solusi terbaik diperoleh berdasarkan nilai *tardiness* terendah..

III. HASIL DAN PEMBAHASAN

Hasil Data Preparation

Data *training* yang digunakan dalam penelitian ini merupakan data historis perusahaan yang mencerminkan aktivitas produksi pada empat tahapan produksi. Tiap tahapan memiliki variabel data yang dibutuhkan untuk pembelajaran mesin dengan jumlah 1000 data pada tiap tahapan. Preprocessing data dilakukan dengan pembersihan data, penanganan nilai kosong, transformasi tipe data, dan normalisasi format atribut. Serta regenerasi data produksi menggunakan Bahasa pemrograman python untuk meningkatkan realism dataset. Regenerasi menghasilkan data untuk merepresentasikan proses produksi yang realistis. Hal ini dilakukan karena data historis cenderung bersifat statis dan belum sepenuhnya mencerminkan variasi kondisi produksi di lapangan. Proses regenerasi dilakukan dengan menggabungkan faktor deterministik dan faktor stokastik. Berikut dataset pada tahapan *sewing* hasil preprocessing sebanyak 1000 data.

Tabel 2 Dataset Sewing

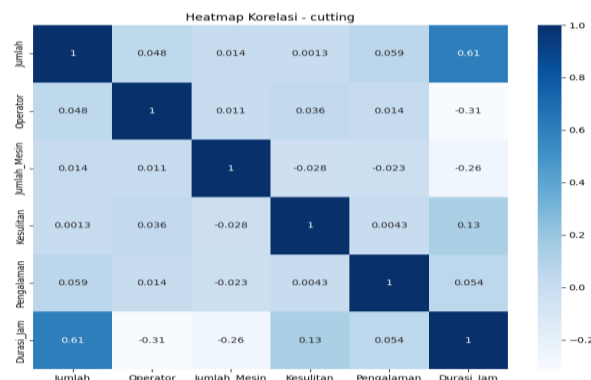
Order_ID	Model	Jenis_Bahan	Jumlah	Operator
ORD001	Kaos	Katun	118	3
ORD002	Topi	Katun	133	3
ORD003	Kaos	Katun	105	3
ORD004	Celana	Denim	68	2
.....				
ORD998	Mitten	Katun	69	5
ORD999	Dress	Satin	144	5
ORD1000	Baby_Set	Baby Terry	44	4

Tabel 3. Lanjutan Dataset Sewing

Kesulitan	Kelompok	Pengalaman	Prioritas	Jumlah_Mesin	Durasi_Jam
2	K3	4	Urgent	11	1.91
1	K1	4	Urgent	11	2.88
3	K3	4	Urgent	6	4.42
1	K2	3	Urgent	6	1.66
.....					
2	K2	2	Urgent	12	1.82
1	K3	4	Normal	6	2.81
2	K2	4	Urgent	5	0.91

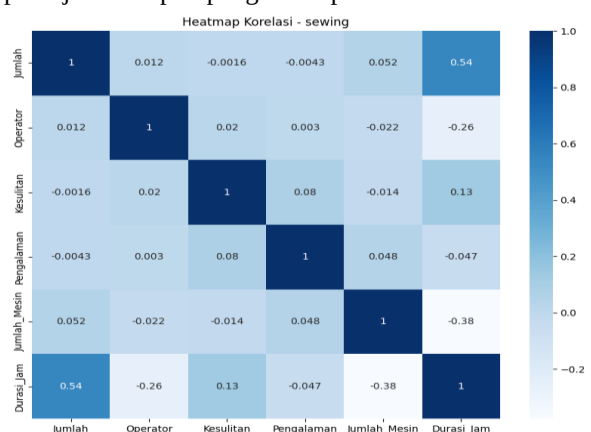
Analisis *heatmap* korelasi menunjukkan bahwa pola hubungan antar variabel pada ketiga tahapan produksi memiliki karakteristik yang konsisten. Pada seluruh tahapan, variabel jumlah produksi memiliki korelasi

positif tinggi terhadap durasi jam yaitu sebesar 0.61 pada proses *cutting*, 0,54 pada *sewing*, dan 0,62 pada *finishing*. Nilai tersebut menunjukkan bahwa semakin volume pekerjaan merupakan faktor dominan yang mempengaruhi lamanya waktu produksi. Semakin banyak unit yang harus diproses, semakin besar pula waktu yang dibutuhkan. Korelasi yang relatif kuat ini mengindikasikan bahwa jumlah produksi merupakan variabel yang harus dipertimbangkan dalam proses produksi



Gambar 2 Heatmap korelasi cutting

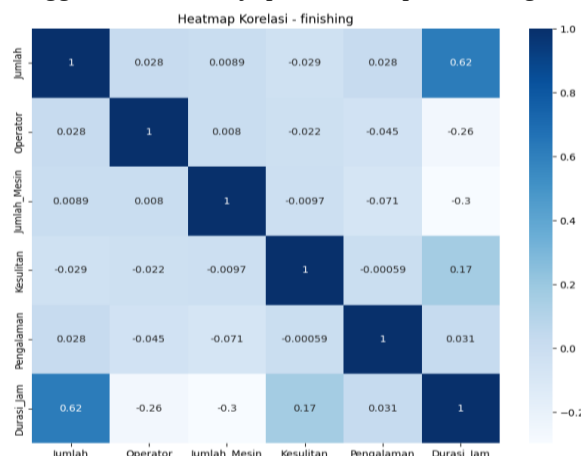
Pada gambar 2 heatmap tahapan *cutting*, variabel Operator dan jumlah_mesin menunjukkan korelasi negatif terhadap durasi masing-masing -0,31 dan -0,26. Hasil ini menunjukkan bahwa peningkatan jumlah operator maupun mesin cenderung menurunkan waktu proses produksi karena kapasitas kerja yang tersedia menjadi lebih besar. Sementara, variabel kesulitan 0,13 dan pengalaman 0,05 memiliki hubungan yang relatif lemah terhadap durasi. Kondisi tersebut mengindikasikan bahwa proses *cutting* lebih dipengaruhi oleh volume pekerjaan dan kapasitas sumber daya dibandingkan tingkat kompleksitas pekerjaan maupun pengalaman operator.



Gambar 3 Heatmap korelasi sewing

Pada gambar 3 heatmap tahapan *sewing*, selain jumlah produksi yang memiliki korelasi positif cukup tinggi terhadap durasi sebesar 0,54, variabel jumlah_mesin memiliki korelasi negatif terhadap durasi yaitu -0,38. Hal ini menunjukkan bahwa kapasitas mesin jahit

merupakan faktor yang sangat menentukan efisiensi proses sewing. Penambahan mesin memungkinkan distribusi pekerjaan yang lebih merata sehingga waktu tunggu dan beban kerja per mesin dapat dikurangi.



Gambar 4 Heatmap korelasi finishing

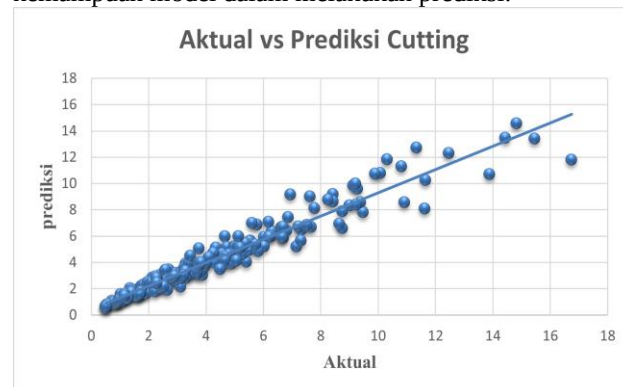
Pada gambar 4 heatmap tahapan *finishing*, korelasi antara jumlah produksi dan durasi mencapai nilai tertinggi dibandingkan tahapan lainnya, yaitu 0,62. Hasil ini menunjukkan bahwa aktivitas *finishing* sangat sensitif terhadap perubahan volume produksi karena setiap produk harus melalui proses pemeriksaan, perapihan dan pengemasan secara individual. Variabel operator -0,26 dan jumlah mesin -0,30 tetap menunjukkan hubungan negatif terhadap durasi, yang mengindikasikan bahwa peningkatan kapasitas sumber daya dapat membantu percepatan penyelesaian pekerjaan. Variabel kesulitan memiliki korelasi positif sebesar 0,17, sedikit lebih tinggi dari tahapan lainnya, yang menunjukkan bahwa kompleksitas produk mulai memberikan pengaruh terhadap durasi produksi.

Secara keseluruhan, *heatmap* korelasi menunjukkan bahwa faktor yang paling berpengaruh terhadap durasi produksi adalah jumlah produksi, diikuti oleh jumlah mesin dan operator. Pola ini mencerminkan karakteristik sistem produksi garmen yang realistis, dimana durasi proses ditentukan oleh besarnya beban kerja serta kapasitas sumber daya yang tersedia. Tidak ditemukan korelasi yang sangat tinggi antar variabel sehingga risiko multikolinearitas relatif rendah. Implikasi dari hasil analisis ini adalah bahwa variabel jumlah produksi, jumlah mesin dan operator menjadi fitur yang paling berkontribusi dalam pembentukan model RF untuk prediksi durasi produksi. Selain itu, pola korelasi yang sesuai dengan pengetahuan domain manufaktur mengindikasikan bahwa dataset simulatif hasil regenerasi berhasil mempertahankan hubungan kausal yang logis dari data historis.

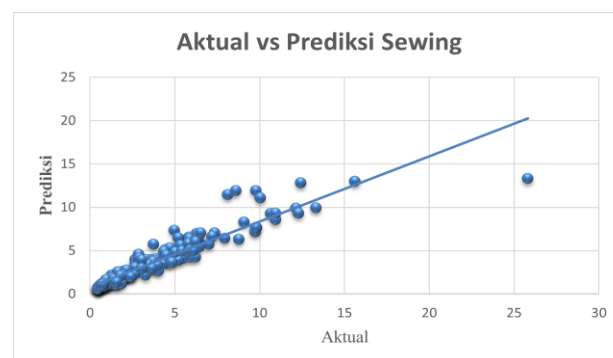
Modeling Prediksi Random Forest

Model *random forest* digunakan untuk memprediksi waktu proses setiap *job*. Proses pelatihan dilakukan dengan membagi data menjadi data latih dan data uji menggunakan perbandingan 80:20. Hasil kinerja pemodelan prediksi random forest divisualisasi menggunakan grafik hubungan antara nilai actual dan nilai prediksi masing masing tahapan. Semakin dekat

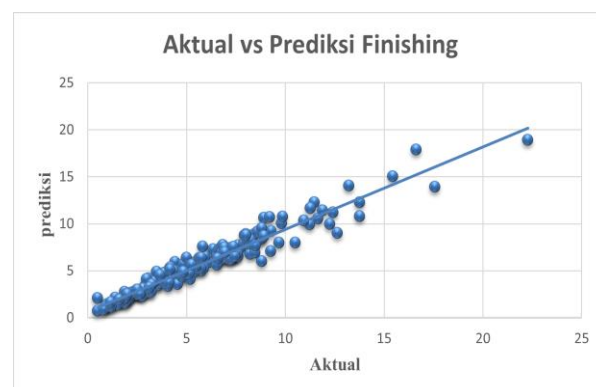
titik data terhadap garis diagonal, maka semakin baik kemampuan model dalam melakukan prediksi.



Gambar 5 Aktual vs prediksi cutting



Gambar 6 Aktual vs prediksi sewing



Gambar 7 Aktual vs prediksi finishing

Pada gambar 5 dan gambar 6 tahapan *cutting* dan *sewing*, Sebagian besar titik data berada disekitar garis diagonal sehingga menunjukkan bahwa *random forest* mampu memprediksi waktu proses *cutting* dan *sewing* dengan cukup baik. Namun beberapa data dengan durasi produksi yang tinggi terlihat adanya penyimpangan yang lebih besar dari garis referensi. Hal ini menunjukkan tingkat kesalahan prediksi meningkat pada pekerjaan dengan kompleksitas atau volume produksi besar. Pada gambar 7, tahapan *finishing* menunjukkan penyebaran data yang relatif rapat menunjukkan bahwa kesalahan prediksi yang dihasilkan lebih kecil dibandingkan tahapan produksi lainnya. Kondisi ini mengindikasikan bahwa model RF memiliki kemampuan prediksi yang baik pada proses *finishing*.

Berdasarkan analisis pada ketiga model, variabel jumlah memiliki nilai *importance* tertinggi dibandingkan variabel lainnya. Pada proses *cutting*, nilai

importance variabel mencapai sekitar 39%, sedangkan pada proses *finishing* mencapai lebih dari 52%. Hal ini menunjukkan bahwa banyaknya unit produk yang diproses merupakan faktor utama mempengaruhi durasi produksi. Variabel model juga memberikan kontribusi cukup besar terhadap waktu produksi. Hal ini disebabkan perbedaan tingkat kompleksitas pengerjaan antar model yang diproduksi. Pada proses *sewing*, variabel jumlah_mesin memiliki pengaruh relatif tinggi, yang menunjukkan bahwa kapasitas mesin menjadi faktor penting dalam menentukan kecepatan proses jahit.

Pada ketiga tahapan produksi, Sebagian besar residual tersebar di sekitar garis nol yang menunjukkan bahwa model tidak mengalami bias sistematis dalam melakukan prediksi. Pada proses *cutting* dan *sewing* penyebaran residual semakin besar pada nilai prediksi yang tinggi, hal ini menunjukkan bahwa kesalahan prediksi cenderung meningkat pada pekerjaan dengan durasi produksi yang lebih Panjang. Sebaliknya, pada proses *finishing* penyebaran residual relatif merata dan tidak menunjukkan pola tertentu. Hal ini mengindikasikan bahwa model memiliki tingkat kestabilan yang lebih baik dalam memprediksi waktu proses *finishing*.

Modeling Optimasi Penjadwalan Algoritma Genetik
Setelah model data diproses menggunakan *random forest*. Kemudian hasil prediksi durasi waktu produksi di optimasi dengan menggunakan Algoritma genetik untuk mencari urutan *job* yang dapat meminimalkan keterlambatan (*tardiness*) dengan tetap mempertahankan efisiensi waktu produksi (*makespan*). GA bekerja dengan merepresentasikan solusi sebagai permutasi urutan *job*. Setiap kromosom merepresentasikan satu alternatif urutan produksi.

Populasi awal dibentuk secara acak dengan ukuran populasi $N = 30$.

$$Pop = X_1, X_2, \dots, X_N$$

Setiap individu dalam populasi merupakan permutasi dari seluruh *job*. Tahap ini menghasilkan variasi solusi awal yang cukup luas agar proses eksplorasi solusi optimal dapat berjalan dengan baik. Evaluasi kromosom menggunakan model *flow shop* untuk menghitung waktu penyelesaian atau *completion time*. Nilai *completion time* tiap *job* yang digunakan untuk menghitung *tardiness*, lalu dicari nilai total *tardiness* dan pembobotan nilai *tardiness*.

Setelah nilai *weighted tardiness* diperoleh, evaluasi kromosom dilanjutkan untuk mencari nilai *makespan*. Nilai *tardiness* dan *makespan* digunakan untuk mencari fungsi *fitness*. Untuk mencari nilai *fitness* dilakukan normalisasi pada nilai *weighted tardiness* dan *makespan* untuk membuat skala yang seimbang dengan pembobotan nilai $\alpha = 0.5$ dan $\beta = 0.5$. Pembobotan ini tidak dipilih secara arbitrer, tetapi mengacu pada konsep *weighted sum multi-objective optimization* yang memberikan bobot sesuai tingkat kepentingan masing-masing tujuan. Dalam penelitian ini pemilihan bobot yang digunakan karena *weighted tardiness* dan *makespan* memiliki tingkat kepentingan yang sama dalam sistem produksi.

Dengan jumlah generasi 100 kali. Algoritma akan berhenti Ketika jumlah generasi maksimum tercapai.

Evaluasi Model Hybrid

A. Evaluasi Random Forest

Hasil evaluasi kinerja model *random forest* dilakukan menggunakan beberapa matrik yaitu *mean absolute error* (MAE) untuk mengukur rata-rata kesalahan antara nilai aktual dan prediksi, *root mean squared error* (RMSE) untuk memberikan penalti lebih besar pada kesalahan yang tinggi dan koefisien determinasi (R^2) untuk mengukur seberapa baik model menjelaskan variabilitas data

Tabel 4 Hasil Evaluasi Model RF

Tahapan	MAE	RMSE	R^2
<i>Cutting</i>	0,407	0,571	0,817
<i>Sewing</i>	0,503	0,633	0,981
<i>finishing</i>	0,454	0,574	0,867

Pada tabel 4 hasil evaluasi model *random forest* tahap *Sewing* mencapai nilai R^2 tertinggi sebesar 0,981, yang berarti model mampu menjelaskan 98,1% variansi durasi pada proses penjahitan. Nilai R^2 yang sangat tinggi ini mengindikasikan bahwa durasi proses *sewing* sangat dapat dijelaskan oleh variabel-variabel *input* yang digunakan, sehingga model RF berhasil menangkap pola dengan sangat baik. Tahap *Cutting* dan *Finishing* juga menunjukkan performa yang memuaskan dengan R^2 masing-masing 0,817 dan 0,867.

Pada proses *finishing* akurasi lebih rendah dibanding dengan tahap *sewing* hal ini dipengaruhi oleh variasi aktivitas *finishing* yang tidak sepenuhnya direpresentasikan oleh variabel input seperti proses pemeriksaan kualitas, perbaikan minor, pelabelan dan pengemasan yang dapat berbeda pada setiap jenis produk. Proses *cutting* menghasilkan nilai paling rendah dari tahap lain hal ini menunjukkan durasi *cutting* memiliki variasi yang lebih tinggi dan dipengaruhi faktor operasional yang belum sepenuhnya dimodelkan seperti penataan pola, pergantian jenis bahan, dll. Faktor-faktor tersebut dapat menyebabkan variasi waktu proses yang sulit diprediksi secara sempurna. Namun, Kedua nilai tersebut mengindikasikan bahwa model cukup andal dalam memprediksi durasi kedua proses tersebut, meskipun masih terdapat sekitar 13–18% variansi yang belum dapat dijelaskan, kemungkinan karena faktor eksternal yang tidak tercakup dalam dataset simulatif.

B. Evaluasi Hybrid Model Random Forest dan Algoritma Genetik

Dilakukan pengujian komputasi dengan metode pembandingan yaitu metode konvensional FCFS dan EDD untuk mengetahui tingkat performa model. Algoritma genetik memiliki karakteristik stokastik dimana hasil tiap *running* sistem dapat berubah dan nilai yang dihasilkan dapat berbeda. Maka dilakukan uji komputasi untuk mengevaluasi signifikansi perbedaan performa antar metode. Eksperimen komputasi dilaksanakan sebanyak 10 kali *running independen*, *population size* 30, dan jumlah generasi 100. Setiap *running* menggunakan set data uji yang sama untuk memastikan

perbandingan yang adil antar metode. Hasil pengujian disajikan dalam tabel berikut:

Tabel 5 Hasil Uji Komputasi Nilai Tardiness Model Hybrid

Run	TardinessF CFS	Tardiness EDD	Tardiness Hybrid RF-GA
1	1839.34	883.20	574.72
2	1839.34	883.20	584.47
3	1839.34	883.20	584.47
4	1839.34	883.20	584.47
5	1839.34	883.20	583.91
6	1839.34	883.20	583.91
7	1839.34	883.20	583.91
8	1839.34	883.20	583.91
9	1839.34	883.20	583.91
10	1839.34	883.20	574.54
mean	1839.34	883.20	582.22

Tabel 6 Hasil Uji Komputasi Nilai Makespan Model Hybrid

Run	MakespanFC FS	Makespan EDD	Makespan Hybrid RF-GA
1	313.85	305.98	299.28
2	313.85	305.98	302.93
3	313.85	305.98	302.93
4	313.85	305.98	302.93
5	313.85	305.98	302.93
6	313.85	305.98	302.93
7	313.85	305.98	302.93
8	313.85	305.98	302.93
9	313.85	305.98	302.93
10	313.85	305.98	299.77
mean	313.85	305.98	302.25

Persamaan untuk Improvement rata-rata nilai *tardiness* dari tiap model :

$$\text{Improvement} = \frac{\text{meanModel}_1 - \text{meanModel}_2}{\text{meanModel}_1} \times 100\% \quad (5)$$

Improvement Hybrid RF-GA terhadap FCFS

$$= \frac{1839.34 - 582.22}{1839.34} \times 100\% = 68.35\%$$

Improvement Hybrid RF-GA terhadap EDD :

$$= \frac{883.20 - 582.22}{883.20} \times 100\% = 34.08\%$$

Improvement EDD terhadap FCFS :

$$= \frac{1839.34 - 883.20}{1839.34} \times 100\% = 51.98\%$$

Persamaan standar deviasi model hybrid :

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}} \quad (6)$$

Standar deviasi *tardiness* hybrid RF-GA :

$$s = \sqrt{\frac{144.71}{9}}$$

$$s = \sqrt{16.08} = 4.01$$

Standar deviasi *makespan* hybrid RF-GA :

$$s = \sqrt{\frac{18.65}{9}}$$

$$s = \sqrt{2.07} = 1.44$$

Persamaan koefisien Variasi untuk menunjukkan kestabilan model hybrid secara statistik :

$$CV = \frac{s}{\text{mean}} \times 100\% \quad (7)$$

Koefisien variasi *tardiness* hybrid RF-GA :

$$CV_{\text{tardiness}} = \frac{4.01}{582.22} \times 100\% = 0.69\%$$

Koefisien variasi *makespan* hybrid RF-GA :

$$CV_{\text{makespan}} = \frac{1.44}{302.25} \times 100\% = 0.48\%$$

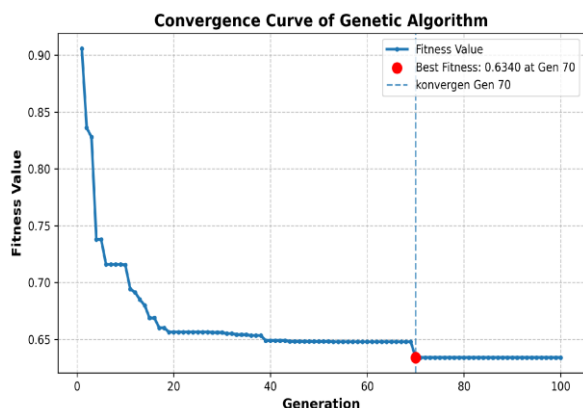
Berdasarkan hasil pengujian nilai *tardiness* pada tabel 5 dan nilai *makespan* pada tabel 6 dapat dijelaskan bahwa dengan pengujian menggunakan data yang sama sebanyak 10 kali *running*, metode FCFS dan EDD termasuk ke dalam algoritma penjadwalan deterministik karena menghasilkan urutan pekerjaan yang tetap untuk data uji yang sama. Metode FCFS menentukan prioritas berdasarkan urutan kedatangan pekerjaan, menghasilkan nilai yang sama pada setiap pengujian dengan nilai rata-rata tertinggi dibanding metode lain pada *tardiness* mencapai 1839.34 jam dan nilai rata-rata *makespan* 313.85 jam.

Penurunan *tardiness* dan *makespan* pada metode EDD dibanding metode FCFS mencapai nilai rata-rata *tardiness* 883.20 jam dan *makespan* 305.98 jam. Persentase perbaikan pada persamaan 5 untuk keterlambatan dari model EDD terhadap FCFS mencapai 51.98 % dan peningkatan *makespan* sebesar 2.51% . Hasil ini menunjukkan bahwa pengurutan berdasarkan *due date* dapat mengurangi keterlambatan signifikan dan menghasilkan aliran pekerjaan yang lebih efisien dibanding urutan kedatangan. Meskipun EDD mampu memberikan peningkatan yang signifikan, namun EDD hanya menggunakan aturan prioritas berdasarkan waktu jatuh tempo terdekat tanpa mempertimbangkan faktor penting lainnya. Akibatnya pekerjaan akan yang memiliki *due date* yang sama diperlakukan identik meskipun memiliki kebutuhan waktu produksi yang sangat berbeda.

Hasil pengujian metode hybrid RF-GA mampu memberikan pengurangan keterlambatan tambahan sebesar 34.08% terhadap EDD atau hampir 301 jam, sepertiga keterlambatan yang masih dapat dieliminasi setelah EDD diterapkan. Metode RF-GA mampu mengurangi keterlambatan 68.35% terhadap FCFS dan mampu meningkatkan nilai *makespan* sebesar 3.70% terhadap metode FCFS. Nilai rata-rata *tardiness* dan *makespan* yang dihasilkan oleh model hybrid secara konsisten lebih rendah dibanding metode konvensional lain sebesar 582,22 jam keterlambatan dan 302,25 jam *makespan*.

Standar deviasi model hybrid RF-GA menggunakan persamaan 6, sebesar 4.01 jam untuk *tardiness* dan 1.44 jam untuk *makespan* menunjukkan

bahwa solusi yang dihasilkan relatif stabil antar pengujian meskipun genetic algorithm bersifat stokastik. Nilai koefisien variasi pada persamaan 7, menunjukan variasi hasil yang rendah berada dibawah 1% pada setiap parameter pengujian model hybrid, hal ini menunjukan bahwa proses pencarian solusi secara konsisten mengarah pada wilayah solusi optimum.



Gambar 8 Kurva konvergensi fitness GA

Gambar 8 menunjukan kurva konvergensi genetic algorithm berdasarkan nilai fitness terbaik pada setiap generasi. Pada generasi awal terjadi penurunan fitness yang dari 0.906 menjadi 0.738, menunjukan bahwa proses seleksi, crossover dan mutasi berhasil memperbaiki kualitas solusi awal secara cepat. Selanjutnya fitness terus menurun secara bertahap hingga mencapai nilai 0.656 pada generasi ke 20. Nilai fitness terbaik sebesar 0.634 diperoleh pada generasi ke 70 dan tidak mengalami perubahan hingga generasi ke 100. Kondisi tersebut menunjukan bahwa algoritma telah mencapai konvergensi dan menemukan solusi optimum. Secara keseluruhan terjadi penurunan fitness sebesar 30.02% dibanding populasi awal. Hasil ini mengindikasikan genetic algorithm mampu melakukan optimasi penjadwalan secara efektif dengan meminimalkan kombinasi nilai tardiness dan makespan.

C. Penentuan Hasil Optimasi Penjadwalan Terbaik Model Hybrid

Dari hasil pengujian model hybrid diperoleh hasil penjadwalan terbaik pada running ke-10 berdasarkan nilai *tardiness* terkecil dengan total keterlambatan 7 job.

Tabel 7 Hasil Penjadwalan Produksi Metode Hybrid

Order_ID	Start	Finish	Deadline
ORD9	30-04-2026 15:22	05-05-2026 08:37	01-05-2026 08:00
ORD17	02-05-2026 08:50	06-05-2026 12:26	01-05-2026 08:00
ORD5	04-05-2026 11:27	08-05-2026 09:33	01-05-2026 08:00
ORD19	06-05-2026 08:27	09-05-2026 14:40	01-05-2026 08:00
.....			
ORD13	04-06-2026 13:48	09-06-2026 08:52	24-06-2026 16:00
ORD3	04-06-2026 11:04	11-06-2026 13:35	04-07-2026 16:00
ORD6	08-06-2026 10:05	13-06-2026 14:55	24-06-2026 16:00

IV. KESIMPULAN

Penelitian ini berhasil menerapkan metode Hybrid Random Forest dan Genetic Algorithm untuk optimasi penjadwalan produksi *flow shop*, di mana Random Forest digunakan untuk memprediksi durasi proses berdasarkan estimasi awal (*time study*) dan kondisi operasional, sedangkan Genetic Algorithm mengoptimalkan urutan job menggunakan fungsi fitness berbasis normalisasi makespan dan tardiness. Model *random forest* menghasilkan prediksi waktu proses yang baik pada tahapan *cutting*, *sewing*, dan *finishing* dengan nilai R^2 masing-masing 0,817, 0,981, dan 0,867. Sementara hasil pengujian model hybrid Random forest dan Genetic Algorithm terbukti menurunkan *tardiness* hingga 68.35% terhadap FCFS dan 34.08% terhadap EDD.

Dengan penurunan keterlambatan yang signifikan, perusahaan dapat meningkatkan ketepatan waktu pengiriman, mengurangi potensi biaya keterlambatan, serta meningkatkan efisiensi operasional perusahaan. Kurva konvergensi *Genetic Algorithm* memperlihatkan penurunan nilai fitness dari 0,906 menjadi 0,634 dan mencapai kondisi stabil pada generasi ke-70, yang menunjukkan kemampuan algoritma dalam menemukan solusi optimum atau mendekati optimum secara efektif. Selain itu, standar deviasi tardiness sebesar 4,01 jam dan makespan sebesar 1,44 jam mengindikasikan bahwa model memiliki tingkat stabilitas dan konsistensi yang tinggi pada pengujian berulang. Temuan ini memberikan implikasi bahwa penerapan pendekatan *hybrid* berbasis *machine learning* dan optimasi metaheuristik dapat menjadi alternatif yang lebih adaptif dan akurat dibandingkan metode penjadwalan konvensional.

Meskipun demikian, penelitian ini masih memiliki keterbatasan pada data yang belum sepenuhnya merepresentasikan kondisi proses produksi dalam skala besar yang dinamis. Untuk itu pengembangan selanjutnya disarankan untuk menggunakan data riil yang lebih lengkap dan sistematis. Serta eksplorasi performa model dibanding dengan metode optimasi lain seperti PSO, *simulated annealing* atau ACO guna memperoleh solusi yang lebih robust dan representatif.

V. REFERENSI

- [1] L. Ramang, R. Budiman, and N. H. Djanggu, "Optimalisasi Proses Produksi Di Konveksi XYZ dengan Line Balancing," *NTEGRATE: Industrial Engineering and Management System*, vol. 9, pp. 419–426, Jun. 2025, [Online]. Available: <https://jurnal.untan.ac.id/index.php/jtinUNTAN/issue/view/2642>
- [2] Sofi Hendra Mukti Wibowo, "Optimasi Penjadwalan No-Wait Flowshop Dengan Sequence-dependent setup Time menggunakan Algoritma Evolusi," Apr. 2025. Accessed: Apr. 26, 2026. [Online]. Available: <https://dspace.uui.ac.id/handle/123456789/59848>
- [3] Z. Zhou, C. Qiu, and Y. Zhang, "A comparative analysis of linear regression, neural networks and random forest regression for predicting air ozone employing soft sensor models," *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-49899-0.
- [4] D. Siregar, L. Sofinah Harahap, and M. F. Alamsyah, "Optimasi Penjadwalan Produksi dengan Jaringan Syaraf Tiruan dan Algoritma Genetika," *Journal of Multidisciplinary*

- Inquiry in Science Technology and Educational Research*, vol. 2, no. 1, 2024, doi: 10.32672/mister.v2i1.2324.
- [5] Nurisa Efida, "USULAN PENJADWALAN PRODUKSI UNTUK MINIMASI TOTAL MAKESPAN MENGGUNAKAN METODE CAMPBELL DUDECK SMITH (CDS) PADA YULIATI WARNO BATIK," May 2024.
- [6] A. Muzamil and M. Nuruddin, "Analisis Penjadwalan Produksi Kaos Dengan Menggunakan Metode PDCA (studi kasus: Al Ghani Konveksi)," *Jurnal Sistem Dan Teknik Industri*, vol. 4, no. 2, Nov. 2023.
- [7] L. Dai *et al.*, "A Novel Production Scheduling Approach Based on Improved Hybrid Genetic Algorithm," *Sustainability (Switzerland)*, vol. 14, no. 18, Sep. 2022, doi: 10.3390/su141811747.
- [8] A. Agusnur, "Optimasi Penjadwalan Produksi dengan Algoritma Genetika," Oct. 2025. [Online]. Available: <http://pustakajurnal.web.id/index.php/jtii>
- [9] S. Hashemi, S. Samadianfard, and A. A. Sadraddini, "Evaluation of Random Forest-Genetic Algorithm Hybrid Model in Estimating Daily Solar Radiation," *Environment and Water Engineering*, vol. 8, no. 3, pp. 636–653, Sep. 2022, doi: 10.22034/jewe.2022.312038.1654.
- [10] M. Norouzi, Z. Gürkaş-Aydın, Ö. C. Turna, M. Y. Yağci, M. A. Aydın, and A. Souri, "A Hybrid Genetic Algorithm-Based Random Forest Model for Intrusion Detection Approach in Internet of Medical Things," *Applied Sciences (Switzerland)*, vol. 13, no. 20, Oct. 2023, doi: 10.3390/app132011145.
- [11] A. Cubukcuoglu, I. Karacan, Z. Ceylan, and S. Bulkan, "Minimizing Makespan in Ordered Flow Shop Scheduling Using a Robust Genetic Algorithm," *Processes*, vol. 13, no. 5, May 2025, doi: 10.3390/pr13051583.
- [12] Y. Y. Zhang, S. H. Chen, Y. W. Wang, C. H. Liao, and C. H. Yu, "A Random Forest-Enhanced Genetic Algorithm for Order Acceptance Scheduling with Past-Sequence-Dependent Setup Times," *Mathematics*, vol. 13, no. 16, Aug. 2025, doi: 10.3390/math13162672.
- [13] Y. Albetian Pangalila, A. Saifudin, H. S. Grace Priscila, and R. Syarial, "Implementasi Algoritma Genetika untuk Optimasi Penjadwalan Produksi pada Industri Manufaktur," *JRIIN: Jurnal Riset Informatika dan Inovasi*, pp. 200–209, 2024, [Online]. Available: <https://jurnalmahasiswa.com/index.php/jriin>
- [14] H. Zhou, W. Cheung, and L. C. Leung, "Minimizing weighted tardiness of job-shop scheduling using a hybrid genetic algorithm," *Eur. J. Oper. Res.*, vol. 194, no. 3, pp. 637–649, May 2009, doi: 10.1016/J.EJOR.2007.10.063.
- [15] F. Bin Shaikat, R. Islam, A. Tabassum Happy, and S. Ahmed Faysal, "Letters in High Energy Physics Optimization of Production Scheduling in Smart Manufacturing Environments Using Machine Learning Algorithms," 2025. [Online]. Available: <https://orcid.org/0009-0004-8396-1874>
- [16] D. Bai *et al.*, "Scheduling a Distributed Permutation Flowshop With Uniform Machines and Release Dates," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 215–227, 2025, doi: 10.1109/TASE.2023.3349167.
- [17] Y. Yuniarthe, R. Fitria Purnomo, R. A. Sari, F. Dirayati, and B. Hartanto, "Application of Bio-Inspired Particle Swarm Optimization Algorithm for Production Scheduling Optimization Corresponding Author," 2025. [Online]. Available: <http://creativecommons.org/licenses/by-sa/4.0/>