

PREDIKSI *BURNOUT* PADA *PROGRAMMER* MENGGUNAKAN TEKNIK PENGENALAN POLA UNTUK IDENTIFIKASI DINI DAN INTERVENSI

Candra Heru Saputra^{*1}, Arief Hermawan², Donny Avianto³

^{1,2,3}Universitas Teknologi Yogyakarta, Yogyakarta
Email: ¹CandraHeru.6230211005@student.uty.ac.id, ²ariefdb@uty.ac.id, ³donny@uty.ac.id
^{*}Penulis Korespondensi

(Naskah masuk: 04 Juni 2024, diterima untuk diterbitkan: 14 Juni 2024)

Abstrak

Burnout atau kelelahan kerja merupakan sebuah fenomena yang sering dihadapi oleh profesional dalam berbagai bidang, termasuk *programmer*. Dampak negatif dari *burnout* mencakup penurunan kesejahteraan individu dan produktivitas kerja. Penelitian ini berkontribusi pada subjek penelitian dengan mengembangkan sebuah model prediktif yang inovatif untuk identifikasi dini dan intervensi *burnout* pada *programmer* menggunakan teknik pengenalan pola. Originalitas penelitian ini terletak pada penerapan teknik pengenalan pola secara khusus pada populasi *programmer*, yang belum banyak dieksplorasi dalam literatur sebelumnya. Data yang digunakan dalam penelitian ini diperoleh dari kuesioner yang mencakup pertanyaan terkait pola kerja, kebiasaan individu, dan indikator *burnout* berdasarkan kriteria *Maslach Burnout Inventory* (MBI). Metodologi yang diterapkan melibatkan pengumpulan dan pra-pemrosesan data, ekstraksi fitur, dan aplikasi algoritma pengenalan pola untuk konstruksi model. Hasil penelitian menunjukkan bahwa model yang dikembangkan mampu mengidentifikasi risiko *burnout* dengan akurasi yang tinggi. Teknik pengenalan pola terbukti efektif dalam menggali pola dan *insight* yang relevan untuk identifikasi dan intervensi *burnout* pada *programmer*, sehingga dapat memberikan kontribusi signifikan dalam pemahaman dan pencegahan *burnout* di kalangan *programmer*. Penelitian ini diharapkan dapat digunakan sebagai referensi dalam praktik dan penelitian lebih lanjut.

Kata kunci: *Burnout, Programmer, Pengenalan Pola, Identifikasi Dini, Intervensi*

PREDICTION OF *BURNOUT* IN PROGRAMMERS USING PATTERN RECOGNITION TECHNIQUES FOR EARLY IDENTIFICATION AND INTERVENTION

Abstract

Burnout is a phenomenon frequently encountered by professionals across various fields, including programmers. The negative impacts of *burnout* include reduced individual well-being and decreased work productivity. This study contributes to the subject by developing an innovative predictive model for early identification and intervention of *burnout* in programmers using pattern recognition techniques. The originality of this research lies in the application of pattern recognition techniques specifically to the programmer population, which has not been extensively explored in previous literature. The data used in this study were obtained from questionnaires that included questions related to work patterns, individual habits, and *burnout* indicators based on the *Maslach Burnout Inventory* (MBI) criteria. The methodology involved data collection and preprocessing, feature extraction, and the application of pattern recognition algorithms for model construction. The results indicate that the developed model is capable of identifying *burnout* risk with high accuracy. Pattern recognition techniques proved effective in uncovering relevant patterns and insights for the identification and intervention of *burnout* in programmers, thereby making a significant contribution to the understanding and prevention of *burnout* among programmers. This study is expected to serve as a reference in both practice and further research.

Keywords: *Burnout, Programmer, Pattern Recognition, Early Identification, Intervention*

1. PENDAHULUAN

Di tengah perkembangan era digital yang cepat, para profesional di sektor teknologi informasi, termasuk para *programmer*, kerap kali dihadapkan pada beban kerja yang berat dan tekanan kerja yang

tinggi (Venkatesh et al., 2020). Kondisi kerja semacam ini dapat menyebabkan *burnout* atau kelelahan kerja, yang memiliki konsekuensi negatif tidak hanya pada kesejahteraan individu tetapi juga pada efisiensi dan hasil kerja mereka (Rybinska et al.,

2022; Tulili et al., 2023). Karena itu, fokus penelitian ini adalah untuk merancang sebuah model yang mampu mengenali indikasi dini dari *burnout* pada *programmer* secara efisien dan tepat.

Kelelahan kerja merupakan fenomena psikologis yang kompleks dengan berbagai dimensi. Berdasarkan analisis literatur yang dilakukan oleh RAND Corporation, *burnout* didefinisikan melalui tiga aspek utama, yaitu kelelahan emosional, proses menjadi tidak peka atau depersonalisasi, dan penurunan perasaan capaian pribadi di antara individu yang bekerja dalam bidang layanan kepada masyarakat (Bouskill et al., 2022). Fenomena ini sering terjadi pada individu yang merasa tidak memiliki kontrol atas lingkungan kerja mereka, yang gagal memenuhi atau tidak yakin dengan harapan kerja, atau yang meragukan nilai intrinsik dari pekerjaan yang mereka lakukan (Shaltout et al., 2023).

Sindrom kelelahan (*burnout*) berlebih di tempat kerja, yang dikenali oleh Organisasi Kesehatan Dunia (WHO) sebagai respons terhadap stres pekerjaan kronis yang tidak dikelola secara efektif, telah menginspirasi pembuatan skala *Korean Version of Burnout Syndrome (KBOSS)* oleh Kim et al. (2021). Skala ini merangkum tiga aspek fundamental yaitu kelelahan, sikap sinis, dan inefektivitas dalam menjalankan tugas profesional, sejalan dengan definisi BOS oleh WHO.

Selain itu, penelitian oleh Guseva Canu et al. (2021) berusaha untuk menyelaraskan definisi *burnout* dalam kerangka kesehatan kerja, dengan menghasilkan definisi yang disetujui oleh panel ahli dari 29 negara. Definisi ini menekankan bahwa *burnout* adalah kondisi kelelahan yang timbul dari eksposur yang berkepanjangan terhadap masalah kerja yang terus-menerus, memberikan perspektif yang memiliki relevansi dan aplikasi secara global.

Menurut Tkachenko & Andrushchenko (2020), kelelahan emosional merupakan tantangan penting dalam industri teknologi, terutama bagi para *programmer*. Kelelahan emosional ini mencakup kelelahan fisik, emosional, dan mental yang muncul dari stres yang berlarut-larut, sering kali karena para *programmer* merasa kewalahan dan tak mampu memenuhi tuntutan yang tidak kunjung selesai. Dalam kondisi seperti ini, mereka dapat kehilangan motivasi dan minat yang awalnya membawa mereka ke dalam profesi tersebut.

Burnout tidak hanya mempengaruhi individu secara psikologis dan kesehatan, tetapi juga berdampak pada perilaku dan aspek organisasional, dengan konsekuensi negatif yang signifikan untuk kedua individu dan organisasi (Edu-valsania et al., 2022). Gejala awalnya termasuk penurunan performa dan energi, iritabilitas yang bertambah, serta merosotnya perasaan pencapaian. Tanpa intervensi yang tepat, keadaan ini dapat berkembang menjadi gangguan kronis yang serius.

Penelitian lain menunjukkan berbagai pendekatan dalam penerapan teknik pembelajaran mesin untuk prediksi dan pencegahan masalah kesehatan. Sebagai contoh, dalam penelitian Jain et al (Jain et al., 2020), model *Logistic Regression* digunakan untuk memprediksi status *churn* pelanggan dengan akurasi 85.2385%. Penelitian ini menyoroti pentingnya pemilihan variabel independen yang tepat untuk meningkatkan akurasi prediksi, yang relevan dalam konteks prediksi *burnout* pada *programmer*.

Demikian pula, dalam penelitian Ambrish G et al (G et al., 2022), model *Logistic Regression* digunakan untuk memprediksi penyakit kardiovaskular dengan akurasi terbaik sebesar 87.10%. Pendekatan ini menggarisbawahi bagaimana analisis data medis dan atribut terkait dapat digunakan untuk mendeteksi kondisi kesehatan yang serius, yang bisa diterapkan untuk mengenali gejala *burnout*.

Penelitian Dritsas & Trigka (2022) menggunakan berbagai teknik pembelajaran mesin untuk memprediksi risiko stroke, menunjukkan bahwa *Logistic Regression* dapat mencapai akurasi 78%. Penelitian ini menyoroti pentingnya pemilihan algoritma yang tepat dan teknik pengolahan data yang efektif untuk meningkatkan akurasi prediksi, yang sangat relevan untuk model prediktif *burnout*.

Kontribusi penelitian ini adalah pada pengembangan model prediktif yang inovatif untuk identifikasi dini *burnout* pada *programmer*, yang memanfaatkan teknik pengenalan pola. Originalitas penelitian ini terletak pada fokus khusus pada populasi *programmer*, yang belum banyak dieksplorasi dalam literatur sebelumnya. Model ini diharapkan mampu mengidentifikasi dan mengintervensi risiko *burnout* secara dini, sehingga tindakan preventif dan mitigasi dapat segera diambil untuk mengurangi dampak negatifnya. Melalui pendekatan ini, penelitian ini berusaha menghasilkan sebuah alat yang praktis dan efisien untuk mendukung kesejahteraan dan produktivitas *programmer* dalam menjalankan tugas dan tanggung jawab profesional mereka.

Tujuan utama dari penelitian ini adalah untuk mengonstruksi sebuah model prediktif yang memanfaatkan teknik pengenalan pola. Model ini diharapkan mampu mengidentifikasi dan mengintervensi risiko *burnout* pada *programmer* secara dini, sehingga tindakan preventif dan mitigasi dapat segera diambil untuk mengurangi dampak negatifnya. Melalui pendekatan ini, penelitian ini berusaha menghasilkan sebuah alat yang praktis dan efisien untuk mendukung kesejahteraan dan produktivitas *programmer* dalam menjalankan tugas dan tanggung jawab profesional mereka.

Fokus penelitian ini adalah pada pengembangan model yang berdasarkan pada data yang dikumpulkan melalui kuesioner. Kuesioner ini dirancang untuk mengumpulkan data terkait dengan pola kerja,

kebiasaan individu, dan indikator lain dari *burnout* sesuai dengan kriteria *Maslach Burnout Inventory* (MBI), sebuah instrumen penelitian yang sudah diakui dan banyak digunakan untuk mengukur tingkat *burnout* (Lin et al., 2022; Maslach & Leiter, 2021; Shoman et al., 2021).

Dengan menggunakan data yang telah dikumpulkan dan diproses, penelitian ini akan menerapkan teknik pengenalan pola untuk menggali *insight* dan pola yang relevan, yang kemudian akan digunakan untuk mengembangkan dan mengoptimalkan model prediktif untuk identifikasi dini *burnout* pada *programmer*.

Penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pemahaman dan pencegahan *burnout* di kalangan *programmer*, dengan menyediakan model dan metodologi yang dapat digunakan sebagai referensi dalam praktik dan penelitian lebih lanjut.

2. METODE PENELITIAN

Dalam penelitian ini, metode yang diadopsi berfokus pada pengumpulan dan analisis data yang diperoleh dari respons kuesioner yang telah diisi oleh para responden. Kuesioner dibagi menjadi dua bagian utama untuk mendapatkan gambaran yang holistik dan mendalam tentang kondisi dan pengalaman responden.

Tabel 1. Daftar Pertanyaan Pola Kerja

No	Pertanyaan	Alias
1	Jenis Kelamin?	X1
2	Berapa usia Anda?	X2
3	Berapa lama Anda bekerja sebagai programmer?	X3
4	Bagaimana kualitas tidur Anda?	X4
5	Berapa jam Anda bekerja setiap hari?	X5
6	Seberapa sering Anda mengambil istirahat selama bekerja?	X6
7	Aktivitas apa yang Anda lakukan selama istirahat?	X7
8	Seberapa sering Anda menghadiri rapat atau diskusi dengan tim?	X8
9	Apakah Anda sering bekerja lembur?	X9
10	Bagaimana pola makan Anda selama hari kerja?	X10
11	Seberapa sering Anda berolahraga atau melakukan aktivitas fisik setiap minggu?	X11
12	Apakah Anda memiliki hobi atau kegiatan di luar pekerjaan yang Anda lakukan secara rutin?	X12

Bagian pertama dari kuesioner terdiri dari 12 pertanyaan yang disusun dengan cermat untuk mengumpulkan data terperinci mengenai pola kerja dan kebiasaan sehari-hari responden. Pertanyaan-pertanyaan ini dirancang untuk menggali informasi mengenai berbagai aspek pengalaman kerja responden, seperti durasi jam kerja, tingkat tekanan kerja, dan faktor-faktor lain yang mungkin

mempengaruhi kondisi kerja dan kesejahteraan responden.

Untuk memfasilitasi analisis dan model statistik, setiap pertanyaan dalam bagian ini dikonversi menjadi sebuah variabel yang diberi nama alias, seperti X1, X2, ..., X12 seperti pada Table 1. Variabel-variabel ini akan dijadikan sebagai variabel independen dalam pengembangan model analitik. Dengan cara ini, setiap variabel yang mewakili jawaban atas pertanyaan tertentu dalam kuesioner, dapat dianalisis untuk menilai pengaruhnya terhadap risiko *burnout* di antara responden.

Tabel 2. Daftar Pertanyaan MBI

No	Pertanyaan
1	Saya merasa sangat lelah secara emosional akibat pekerjaan saya
2	Saya mudah memahami apa yang dirasakan rekan kerja saya
3	Saya merasa berinteraksi dengan beberapa rekan kerja seperti berinteraksi dengan benda mati
4	Saya merasa tenaga saya habis di setiap akhir jam kerja
5	Saya merasa sangat lelah ketika bangun di pagi hari dan harus menjalankan tugas di kantor
6	Saya bekerja secara efektif dalam menyelesaikan tugas yang saya kerjakan
7	Saya merasa memberi pengaruh positif terhadap hidup orang lain melalui pekerjaan saya
8	Sejak bekerja sebagai karyawan saya merasa semakin tidak memperhatikan perasaan orang lain
9	Saya khawatir pekerjaan ini membuat saya menjadi pribadi yang keras secara emosional
10	Bekerja dengan orang-orang sepanjang hari benar-benar merupakan tekanan bagi saya
11	Saya merasa sangat bersemangat
12	Saya tidak terlalu peduli dengan apa yang terjadi pada rekan kerja saya
13	Saya merasa teramat letih dalam menyelesaikan tugas rutin saya
14	Saya merasa frustrasi dengan pekerjaan saya sebagai karyawan
15	Saya pandai membuat suasana yang nyaman bersama rekan kerja saya
16	Saya merasa rekan kerja saya menyalahkan saya atas sebagian masalah mereka
17	Saya merasa bekerja terlalu keras dalam menyelesaikan tugas rutin saya di kantor
18	Saya merasa sangat bahagia setelah menyelesaikan tugas rutin saya
19	Bekerja dengan orang lain secara langsung memberikan stres yang berlebihan pada saya
20	Saya telah mencapai banyak hal yang bermanfaat dalam pekerjaan ini
21	Saya merasa sedang di ujung kemampuan saya
22	Dalam bekerja, saya mengatasi masalah emosi dengan sangat tenang

Bagian kedua dari kuesioner ini mengintegrasikan 22 pertanyaan yang bersumber dari *Maslach Burnout Inventory* (MBI), suatu alat ukur yang terkenal dan sering digunakan untuk menilai tingkat kelelahan emosional atau *burnout* pada individu dalam lingkungan kerjanya sebagaimana terdapat pada Tabel 2.

Tabel 3. *Dataset* respons kuesioner

Responden	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	0	1	3	3	2	3	3	4	1	1	3	1	1
2	1	3	5	3	2	3	4	4	2	1	4	0	1
3	1	3	3	1	2	4	2	4	1	4	3	0	0
4	0	3	2	2	4	2	4	1	2	1	4	1	1
5	1	1	3	2	3	2	2	1	2	1	3	1	0
6	0	3	1	1	4	3	1	1	2	2	3	1	1
7	1	3	1	3	3	3	2	4	1	2	1	0	0
...
114	1	4	2	1	3	1	2	4	3	1	3	1	1

Analisis terhadap jawaban dari kuesioner MBI ini dilakukan dengan mempertimbangkan tiga aspek utama, yaitu Kelelahan Emosional (*Emotional Exhaustion* atau EE), Depersonalisasi (*Depersonalization* atau DP), dan Pencapaian Pribadi (*Personal Accomplishment* atau PA).

Kelelahan Emosional (EE) mengevaluasi tingkat kelelahan emosional dan fisik yang dirasakan, menjadi indikator kunci dari burnout. Skor untuk dimensi ini dihitung dengan menambahkan nilai dari pertanyaan-pertanyaan terkait. Depersonalisasi (DP) menilai sejauh mana seseorang mengalami pemisahan emosional atau distansi dalam berinteraksi dengan rekan kerja atau klien, dengan skor yang dihitung serupa. Sementara Pencapaian Pribadi (PA) mengukur tingkat efikasi profesional dan keberhasilan seseorang dalam berinteraksi dan bekerja dengan orang lain, dengan skor yang dihasilkan dari penjumlahan nilai pertanyaan terkait.

Setelah nilai untuk setiap dimensi dihitung, responden kemudian dikategorikan mengalami *burnout* atau tidak berdasarkan nilai ambang tertentu dari dimensi EE, DP, dan PA. Kriteria penilaian adalah sebagai berikut:

- responden dengan skor $EE \geq 27$
- responden dengan skor $DP \geq 10$
- responden dengan skor $PA \leq 33$

Jika seorang responden memenuhi semua kriteria di atas, maka mereka dikategorikan mengalami *burnout*.

Berdasarkan analisis dan perhitungan nilai MBI, status *burnout* kemudian diubah menjadi variabel biner, di mana "1" mewakili *burnout* dan "0" mewakili *non-burnout*. Variabel biner ini, yang diwakili dengan Y, kemudian digunakan sebagai variabel dependen dalam model analisis dan prediksi yang dikembangkan dalam penelitian ini seperti terlihat pada Tabel 3.

Untuk membangun model analisis dan prediksi yang efektif, penelitian ini menggunakan *Logistic Regression* sebagai metode utama. *Logistic Regression* merupakan metode yang sering digunakan untuk menghasilkan model prediksi yang memberikan probabilitas terjadinya suatu peristiwa (Sakinah et al., 2020) dan terbukti memiliki tingkat akurasi yang tinggi dalam melakukan prediksi (Agustina et al., 2023).

Selain itu *Logistic Regression* sangat cocok untuk menangani variabel dependen biner (Arabameri et al., 2019), seperti dalam kasus variabel Y yang mewakili status *burnout* (1) dan *non-burnout* (0). *Logistic Regression* juga memungkinkan peneliti untuk memperkirakan probabilitas terjadinya suatu peristiwa (*burnout*) berdasarkan kombinasi dari variabel independen (X1, X2, ..., X12).

Dengan menggunakan metodologi ini, penelitian ini bertujuan untuk menghasilkan model prediktif yang andal dan akurat yang dapat digunakan untuk mengidentifikasi risiko *burnout* pada individu berdasarkan respons mereka terhadap kuesioner.

3. HASIL DAN PEMBAHASAN

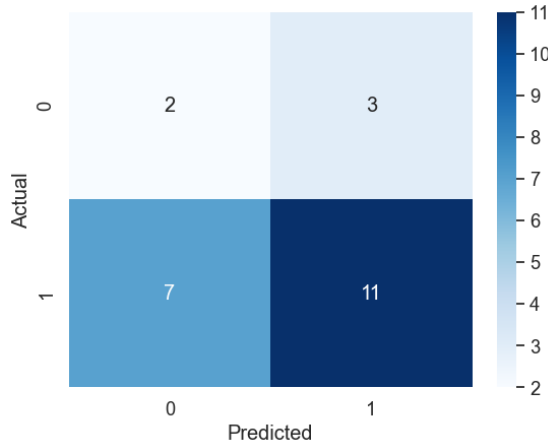
3.1. Hasil dan Pembahasan

Sebagai langkah awal dalam mempersiapkan data untuk analisis, *dataset* tersebut dibagi menjadi dua bagian untuk memastikan bahwa model dapat dilatih dan diuji dengan cara yang paling optimal. Sebanyak 80% dari total data, yang setara dengan sekitar 91 responden, dialokasikan untuk pelatihan model. Bagian ini digunakan untuk "mengajarkan" model tentang karakteristik data dan pola yang ada dalam data. Sementara itu, sisanya, yaitu 20% atau sekitar 23 responden, digunakan sebagai data pengujian. Data pengujian ini memungkinkan kita untuk mengevaluasi sejauh mana kinerja model dalam situasi nyata dan memastikan bahwa model dapat membuat prediksi yang akurat dan andal pada data yang belum pernah dilihat. Rasio pembagian 80:20, di mana 80% data digunakan untuk pelatihan dan 20% untuk pengujian, merupakan pendekatan yang sering digunakan. Justifikasi untuk pembagian ini sering kali dikaitkan dengan prinsip *Pareto* (Joseph, 2022).

Dengan kombinasi *dataset* yang telah disiapkan dengan baik dan pemilihan model *Logistic Regression*, penelitian ini bertujuan untuk mendapatkan wawasan mendalam tentang faktor-faktor yang berkontribusi terhadap *burnout* dan bagaimana kita dapat memprediksinya dengan akurasi yang tinggi.

	precision	recall	f1-score	support
0	0.00	0.00	0.00	5
1	0.75	0.83	0.79	18
accuracy			0.65	23
macro avg	0.38	0.42	0.39	23
weighted avg	0.59	0.65	0.62	23

Gambar 1. Grafik kinerja model analisis awal



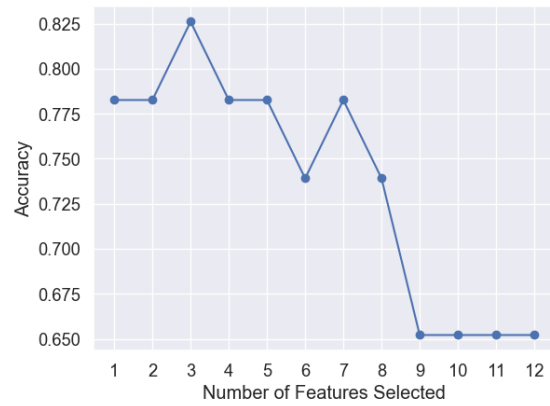
Gambar 2. Confusion Matrix

Hasil analisis awal yang dilakukan dengan menggunakan data uji menunjukkan bahwa model ini mencapai tingkat akurasi sebesar 65%, sebagaimana ditunjukkan pada grafik kinerja Gambar 1 dan grafik *Confusion Matrik* Gambar 2. Nilai akurasi ini mencerminkan kemampuan model untuk mengidentifikasi dan memprediksi kemungkinan seseorang mengalami *burnout* dengan cukup baik, tetapi juga menunjukkan bahwa masih ada ruang untuk melakukan peningkatan dan optimisasi lebih lanjut. Akurasi hanya 65% ini disebabkan oleh adanya beberapa fitur yang tidak memberikan kontribusi signifikan terhadap prediksi *burnout*, sehingga menambah *noise* dalam data dan mengurangi efektivitas model. Beberapa contoh fitur yang mungkin menambah *noise* dan tidak memberikan kontribusi signifikan adalah jenis kelamin (X1), yang kemungkinan tidak memiliki korelasi yang kuat dengan *burnout*, frekuensi menghadiri rapat atau diskusi dengan tim (X8), yang mungkin tidak memiliki korelasi langsung dengan tingkat *burnout*, serta hobi atau kegiatan di luar pekerjaan yang dilakukan secara rutin (X12), yang mungkin tidak relevan atau memiliki hubungan yang lemah dengan *burnout* dalam konteks pekerjaan.

Dalam upaya untuk meningkatkan kinerja model, proses optimisasi dilakukan melalui *feature selection*. Tujuannya adalah untuk menentukan fitur-fitur mana yang paling berkontribusi dan memiliki pengaruh signifikan terhadap kemampuan prediksi model. *Feature selection* membantu dalam mengidentifikasi dan menghilangkan fitur yang kurang relevan atau tidak berkontribusi secara signifikan terhadap model, sehingga model dapat lebih fokus pada fitur-fitur yang memiliki informasi

paling berharga dan relevan untuk memprediksi *burnout*.

Sebagai bagian dari proses optimisasi ini, sebuah grafik dikembangkan untuk menunjukkan hubungan antara jumlah fitur yang dipilih dan akurasi model *Logistic Regression* sebagaimana dapat dilihat pada Gambar 3.



Gambar 3. Grafik akurasi seleksi fitur

Menariknya, dari 12 fitur yang tersedia, model mencapai akurasi tertinggi ketika hanya menggunakan tiga fitur terpilih. Hal ini menegaskan bahwa pemilihan tiga fitur krusial ini dapat meningkatkan efisiensi model tanpa mengorbankan akurasi prediksi.

Tabel 4. Nilai akurasi berdasarkan kombinasi fitur

Nomor	Features	Accuracy
1	X4, X7, X10	0.826086957
2	X2, X7, X10	0.826086957
3	X5, X7, X10	0.826086957
4	X4, X7, X9	0.782608696
5	X4, X5, X10	0.782608696
6	X4, X5, X11	0.782608696
...
220	X2, X10, X11	0.739130435

Selain itu, analisis lebih lanjut menunjukkan bahwa ada kombinasi fitur tertentu yang memberikan hasil prediksi yang lebih akurat. Tiga kombinasi fitur yaitu (X4, X7, X10), (X2, X7, X10), dan (X5, X7, X10) semuanya mencapai akurasi prediksi sebesar 82,6086957% seperti yang terlihat pada Tabel 4. Penemuan ini menyoroti bahwa ketiga kombinasi fitur tersebut memegang peranan krusial dalam meningkatkan performa model. Fitur-fitur ini lebih relevan dan signifikan dalam memprediksi *burnout*, yang menjelaskan peningkatan akurasi menjadi 82,6086957%. Dengan mengeliminasi fitur-fitur yang tidak relevan dan fokus pada fitur yang memiliki kontribusi signifikan, model dapat lebih efektif dalam memprediksi *burnout*, sebagaimana dibuktikan dengan peningkatan akurasi setelah optimisasi. Dengan demikian, kehadiran fitur-fitur ini dalam

model memungkinkan untuk memprediksi burnout dengan akurasi yang lebih tinggi.

Sebelum melakukan *feature selection*, model ini menunjukkan akurasi sebesar 65%. Namun, setelah proses *feature selection*, di mana fitur-fitur yang kurang relevan disingkirkan, akurasi model meningkat menjadi 83%. Hal ini menunjukkan bahwa pengurangan fitur yang tidak esensial dapat membantu model fokus pada informasi yang lebih krusial dan relevan, sehingga meningkatkan performa dan akurasi dalam memprediksi burnout. Grafik hasil kinerja dari model yang telah dilakukan *feature selection* dapat dilihat pada Gambar 4.

	precision	recall	f1-score	support
0	1.00	0.20	0.33	5
1	0.82	1.00	0.90	18
accuracy			0.83	23
macro avg	0.91	0.60	0.62	23
weighted avg	0.86	0.83	0.78	23

Gambar 4. Grafik kinerja model dengan *feature selection*

3.2. Analisis Perbandingan Model

Penulis juga membandingkan kinerja model *Logistic Regression* dengan tiga model pembelajaran mesin lainnya, yaitu *Random Forest*, SVM, dan KNN, menggunakan konfigurasi standar. Hasil perbandingan kinerja model-model tersebut ditampilkan pada Tabel 5 berikut.

Tabel 5. Nilai perbandingan akurasi dan AUC

Model	Akurasi	AUC
Logistic Regression	0.826087	0.433333
Random Forest	0.608696	0.444444
SVM	0.739130	0.566667
KNN	0.695652	0.505556

Logistic Regression menunjukkan akurasi tertinggi sebesar 82,61%, namun nilai AUC-nya hanya 0,433. Meskipun model ini sering memberikan prediksi yang benar, kemampuan diskriminatifnya dalam membedakan antara kelas positif dan negatif terbukti rendah. *Random Forest* menunjukkan akurasi sebesar 60,87% dan AUC sebesar 0,444, yang mengindikasikan bahwa model ini kurang baik dalam mengklasifikasikan data dengan benar. Model SVM menampilkan keseimbangan yang lebih baik dengan akurasi 73,91% dan AUC tertinggi di antara semua model, yaitu 0,567, menunjukkan kemampuan diskriminatif yang moderat. KNN memiliki akurasi 69,57% dan AUC 0,506, yang menunjukkan performa yang hampir sama dengan tebakan acak.

Perbedaan akurasi ini menunjukkan variasi kemampuan masing-masing model untuk mengklasifikasikan data dengan benar. Jumlah *dataset* yang kecil, yaitu 114 data, dapat mempengaruhi hasil secara signifikan. *Dataset* kecil dapat menyebabkan *overfitting*, variabilitas tinggi, kurangnya representasi, dan evaluasi yang tidak

stabil. Meskipun *Logistic Regression* menunjukkan akurasi yang tinggi, SVM mungkin menjadi pilihan yang lebih baik karena memiliki kemampuan yang lebih seimbang dalam membedakan antara kelas positif dan negatif.

SVM dikenal cukup efektif untuk bekerja dengan *dataset* yang kecil karena mampu menemukan *hyperplane* yang memaksimalkan *margin* antara kelas-kelas, sehingga lebih baik dalam menangani masalah *overfitting* yang umum terjadi pada *dataset* kecil. Selain itu, SVM dapat memberikan kinerja yang baik dalam situasi di mana data tidak linier dengan menggunakan *kernel trick*.

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Dalam penelitian ini, sebuah model berbasis *Logistic Regression* berhasil dikembangkan dengan tujuan untuk memprediksi kemungkinan terjadinya burnout di kalangan *programmer*. Model ini menunjukkan efektivitas yang cukup baik, di mana telah berhasil mencapai akurasi yang memuaskan dalam membuat prediksi, meskipun terdapat beberapa aspek yang masih bisa dioptimalkan untuk peningkatan performa lebih lanjut.

Hasil analisis menunjukkan bahwa *Logistic Regression* memiliki akurasi tertinggi sebesar 82,61%, namun nilai AUC-nya hanya 0,433, menunjukkan kemampuan diskriminatif yang rendah. Model SVM menunjukkan keseimbangan yang lebih baik antara akurasi dan AUC, dengan akurasi 73,91% dan AUC tertinggi 0,567. Model KNN dan *Random Forest* menunjukkan performa yang lebih rendah dalam hal akurasi dan AUC.

Salah satu aspek kunci dalam peningkatan kinerja model adalah penggunaan teknik *feature selection* atau pemilihan fitur. Teknik ini sangat krusial karena berkontribusi signifikan dalam meningkatkan akurasi model. Dengan menerapkan teknik ini, model menjadi lebih fokus pada fitur-fitur yang paling berpengaruh dan mengabaikan fitur-fitur yang kurang relevan atau tidak memberikan kontribusi signifikan. Dalam hal ini, dari 12 fitur awal, hanya tiga fitur krusial yang dipilih untuk digunakan dalam model, yang menyebabkan peningkatan akurasi prediksi.

Analisis lebih lanjut terhadap kombinasi fitur menunjukkan bahwa ada beberapa kombinasi fitur tertentu yang sangat efektif. Kombinasi fitur seperti X4, X7, dan X10; X2, X7, dan X10; serta X5, X7, dan X10, semuanya menunjukkan performa yang sangat baik dalam model, memungkinkan model untuk mencapai akurasi prediksi sebesar 82,61%.

Secara keseluruhan, penelitian ini menghasilkan model prediktif yang dapat digunakan untuk mengidentifikasi risiko burnout di kalangan *programmer* dengan akurasi yang cukup tinggi. Penggunaan teknik *feature selection* terbukti efektif dalam meningkatkan performa model.

4.2. Saran

Berikut adalah beberapa saran yang dapat ditindaklanjuti untuk penelitian dan pengembangan lebih lanjut dalam memprediksi dan mengintervensi *burnout* pada programmer:

- a. Optimisasi Model
Meskipun telah dilakukan optimisasi model melalui teknik *feature selection*, masih ada peluang untuk meningkatkan kinerja model. Hal ini dapat dicapai melalui teknik seperti *hyperparameter tuning* atau eksplorasi algoritma *machine learning* lain seperti *Random Forest*, SVM, dan KNN untuk menemukan pendekatan yang mungkin lebih efektif.
- b. Ekspansi Dataset
Dataset yang digunakan dalam penelitian ini terbatas hanya 114 responden. Oleh karena itu, disarankan untuk mengexpand dataset dengan jumlah responden yang lebih banyak untuk meningkatkan keandalan dan kemampuan generalisasi model, sehingga membuat model lebih *robust* dan aplikatif dalam berbagai kondisi.
- c. Analisis Fitur Lebih Lanjut
Saran lain adalah melakukan analisis lebih lanjut terhadap fitur-fitur yang telah dipilih. Tujuannya adalah untuk mendapatkan pemahaman yang lebih mendalam tentang bagaimana masing-masing fitur berkontribusi terhadap risiko *burnout*, dan untuk mengidentifikasi faktor atau variabel mana yang paling kritis dalam menentukan risiko tersebut.
- d. Penerapan Praktis
Model yang telah dikembangkan dapat diuji dan diterapkan lebih lanjut dalam konteks praktis untuk membantu dalam identifikasi dini dan intervensi *burnout* pada programmer. Ini dapat melibatkan pengembangan aplikasi atau sistem pendukung keputusan yang dapat digunakan oleh organisasi dan profesional dalam mengidentifikasi dan mengatasi *burnout*.
- e. Studi Lanjutan
Mengingat relevansi dan urgensi masalah *burnout* dalam profesi *programming*, disarankan untuk melakukan penelitian lebih lanjut. Studi lanjutan dapat mencakup integrasi metode atau pendekatan baru seperti *deep learning* atau model *ensemble* untuk lebih memahami dan mengatasi *burnout* dalam konteks ini.

Setiap saran ini bertujuan untuk memberikan panduan untuk perbaikan dan pengembangan lebih lanjut dari model dan metode yang digunakan dalam penelitian ini, serta untuk penelitian lebih lanjut pada topik ini.

DAFTAR PUSTAKA

AGUSTINA, N., IHSAN, C. N., TINGGI, S., BANDUNG, T., Riset, B., NASIONAL, I., &

KORESPONDENSI, P. 2023. *Pendekatan Ensemble Untuk Analisis Sentimen Covid19 Menggunakan Pengklasifikasi Soft Voting An Ensemble Approach For Covid19 Sentiment Analysis Using Soft Voting Classifier*. 10(2), 263–270.

<https://doi.org/10.25126/jtiik.2023106215>

ARABAMERI, A., PRADHAN, B., & LOMBARDO, L. 2019. Comparative assessment using boosted regression trees, binary logistic regression, frequency ratio and numerical risk factor for gully erosion susceptibility modelling. *CATENA*, 183, 104223.

<https://doi.org/10.1016/J.CATENA.2019.104223>

BOUSKILL, K. E., DANZ, M., MEREDITH, L. S., CHEN, C., CHANG, J., BAXI, S. M., HUYNH, D., AL-IBRAHIM, H., MOTALA, A., LARKIN, J., AKINNIRANYE, O., & HEMPEL, S. 2022. *Burnout Definition, Prevalence, Risk Factors, Prevention, and Interventions Literature Reviews Research Report*. www.rand.org/about/principles.

DRITSAS, E., & TRIGKA, M. 2022. Stroke Risk Prediction with Machine Learning Techniques. *Sensors*, 22(13). <https://doi.org/10.3390/s22134670>

EDÚ-VALSANIA, S., LAGUÍA, A., & MORIANO, J. A. 2022. Burnout: A Review of Theory and Measurement. In *International Journal of Environmental Research and Public Health* (Vol. 19, Issue 3). MDPI. <https://doi.org/10.3390/ijerph19031780>

G, A., GANESH, B., GANESH, A., SRINIVAS, C., DHANRAJ, & MENSINKAL, K. 2022. Logistic regression technique for prediction of cardiovascular disease. *Global Transitions Proceedings*, 3(1), 127–130. <https://doi.org/10.1016/j.glt.2022.04.008>

GUSEVA CANU, I., MARCA, S. C., DELL'ORO, F., BALÁZS, Á., BERGAMASCHI, E., BESSE, C., BIANCHI, R., BISLIMOVSKA, J., BJELAJAC, A. K., BUGGE, M., BUSNEAG, C. I., ÇAĞLAYAN, Ç., CERNITANU, M., PEREIRA, C. C., HAFNER, N. D., DROZ, N., EGLITE, M., GODDERIS, L., GÜNDEL, H., ... WAHLEN, A. 2021. Harmonized definition of occupational burnout: A systematic review, semantic analysis, and Delphi consensus in 29 countries. *Scandinavian Journal of Work, Environment and Health*, 47(2), 95–107. <https://doi.org/10.5271/sjweh.3935>

JAIN, H., KHUNTETA, A., & SRIVASTAVA, S. 2020. Churn Prediction in Telecommunication using Logistic Regression and Logit Boost. *Procedia Computer Science*, 167, 101–112. <https://doi.org/10.1016/j.procs.2020.03.187>

- JOSEPH, V. R. 2022. Optimal ratio for data splitting. *Statistical Analysis and Data Mining*, 15(4), 531–538. <https://doi.org/10.1002/sam.11583>
- KIM, H. D., PARK, S. G., KIM, W. H., MIN, K. B., MIN, J. Y., & HWANG, S. H. 2021. Development of Korean Version Burnout Syndrome Scale (KBOSS) Using WHO's Definition of Burnout Syndrome. *Safety and Health at Work*, 12(4), 522–529. <https://doi.org/10.1016/j.shaw.2021.08.001>
- LIN, C. Y., ALIMORADI, Z., GRIFFITHS, M. D., & PAKPOUR, A. H. 2022. Psychometric properties of the Maslach Burnout Inventory for Medical Personnel (MBI-HSS-MP). *Heliyon*, 8(2). <https://doi.org/10.1016/j.heliyon.2022.e08868>
- MASLACH, C., & LEITER, M. P. 2021. *How to Measure Burnout Accurately and Ethically*.
- RYBINSKA, Y., LOSHENKO, O., KYRYLENKO, T., KONDRATIEVA, V., SERBOVA, O., & STEBAIEVA, O. 2022. Comprehensive Psychological Analysis of The Features of Emotional Burnout Among IT Specialists: The Ukrainian Labor Market. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 13(2), 273–289. <https://doi.org/10.18662/brain/13.2/343>
- SAKINAH, N., BADRIYAH, T., SYARIF, I., & KORESPONDENSI, P. 2020. *Analisis Kinerja Algoritma Mesin Pembelajaran Untuk Klasifikasi Penyakit Stroke Menggunakan Citra Ct Scan Performance Analysis Machine Learning Algorithms For Classification Of Stroke Using Ct Scan Images*. <https://doi.org/10.25126/jtiik.202073482>
- SHALTOUT, A. E., MOHAMED, M. A., IBRAHIM, N. M., & ELDAHSHAN, N. A. 2023. Prevalence of Burnout Syndrome among Working Physicians in Family Health Centres and Units in Port Said Governorate. *Asian Journal of Medicine and Health*, 21(9), 25–43. <https://doi.org/10.9734/ajmah/2023/v21i9853>
- SHOMAN, Y., MARCA, S. C., BIANCHI, R., GODDERIS, L., VAN DER MOLEN, H. F., & GUSEVA CANU, I. 2021. Psychometric properties of burnout measures: A systematic review. In *Epidemiology and Psychiatric Sciences*. Cambridge University Press. <https://doi.org/10.1017/S2045796020001134>
- TKACHENKO, K., & ANDRUSHCHENKO, M. 2020. *How to prevent emotional burnout for programmers*.
- TULILI, T. R., CAPILUPPI, A., & RASTOGI, A. 2023. Burnout in software engineering: A systematic mapping study. In *Information and Software Technology* (Vol. 155). Elsevier B.V. <https://doi.org/10.1016/j.infsof.2022.107116>
- VENKATESH, V., THONG, J. Y. L., CHAN, F. K. Y., HOEHLE, H., & SPOHRER, K. 2020. How agile software development methods reduce work exhaustion: Insights on role perceptions and organizational skills. *Information Systems Journal*, 30(4), 733–761. <https://doi.org/10.1111/isj.12282>