

# Penerapan Machine Learning Untuk Prediksi Gangguan Kesehatan Mental pada Pekerja di Bidang Teknologi

Satriani Satriani<sup>1</sup>, Nayla Dwi Rianti Putri<sup>2\*</sup>, Marsha Selma Rahim<sup>3</sup>, Rina Rina<sup>4</sup>,  
<sup>1,2,3</sup>Program Studi Teknik Informatika, Universitas Muslim Indonesia, Kota Makassar

\*Email Korespondensi: [13020230154@student.umi.ac.id](mailto:13020230154@student.umi.ac.id)

Riwayat Artikel: Diterima: 20/08/2025; Direvisi: 15/09/2025; Disetujui: 25/10/2025

## ABSTRAK

Perkembangan pesat industri teknologi membawa konsekuensi berupa meningkatnya tekanan kerja, jam kerja panjang, serta tuntutan inovasi berkelanjutan yang berpotensi memicu gangguan kesehatan mental pada pekerja. Penelitian ini bertujuan menganalisis hubungan antara dukungan perusahaan dan risiko gangguan kesehatan mental pada pekerja sektor teknologi dengan pendekatan *machine learning*. Dataset yang digunakan adalah *Mental Health in Tech Survey* dari *Open Sourcing Mental Illness* (OSMI) yang terdiri dari 1.259 responden dan 28 variabel terkait kondisi mental, dukungan organisasi, serta interferensi kerja. Tahapan penelitian meliputi *preprocessing data* (pembersihan, encoding, normalisasi, dan penyeimbangan kelas), klasifikasi menggunakan algoritma C4.5 untuk memperoleh aturan keputusan yang interpretabel, serta analisis *clustering* K-Means untuk mengidentifikasi pola kelompok risiko. Hasil analisis menunjukkan bahwa pekerja dengan dukungan perusahaan tinggi seperti ketersediaan benefit kesehatan mental, dukungan supervisor, dan program kesejahteraan cenderung memiliki tingkat interferensi kerja dan kebutuhan treatment yang lebih rendah. Sebaliknya, dukungan yang rendah atau tidak jelas berkorelasi dengan tingkat stres dan gangguan mental yang lebih tinggi. Penelitian ini menegaskan bahwa dukungan organisasi berperan sebagai faktor protektif penting dalam menjaga kesejahteraan psikologis pekerja teknologi. Temuan ini diharapkan menjadi dasar bagi perusahaan dalam merancang kebijakan kesehatan mental yang lebih inklusif dan preventif guna meningkatkan produktivitas dan keberlanjutan organisasi.

**Kata Kunci:** burnout; dukungan Perusahaan; kesehatan mental; machine learning; pekerja teknologi

## 1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat telah membawa perubahan signifikan dalam dunia kerja, khususnya pada sektor teknologi. Pekerja teknologi dituntut untuk bekerja secara cepat, adaptif, dan inovatif dalam menghadapi dinamika industri yang kompetitif. Tuntutan tersebut sering kali disertai dengan beban kerja yang tinggi, tekanan waktu (*deadline*), jam kerja yang panjang, serta ekspektasi kinerja yang terus meningkat. Kondisi ini menjadikan pekerja teknologi sebagai kelompok yang rentan mengalami gangguan kesehatan mental, seperti stres kerja, kecemasan, burnout, hingga depresi [2], [3].

Gangguan kesehatan mental pada pekerja tidak hanya berdampak pada kesejahteraan individu, tetapi juga memengaruhi produktivitas, kualitas kerja, dan keberlangsungan organisasi. Pekerja yang mengalami stres berkepanjangan cenderung mengalami penurunan konsentrasi, motivasi kerja, serta kepuasan kerja. Dalam jangka panjang, kondisi tersebut dapat meningkatkan angka absensi, *turnover* karyawan, dan menurunnya kinerja organisasi secara keseluruhan. Oleh karena itu, isu kesehatan mental di tempat kerja menjadi perhatian penting bagi perusahaan, khususnya pada sektor teknologi yang memiliki karakteristik kerja dengan tekanan tinggi [1], [4].

Salah satu faktor yang diyakini berperan penting dalam menjaga kesehatan mental pekerja adalah dukungan perusahaan. Dukungan perusahaan dapat diwujudkan dalam berbagai bentuk, seperti dukungan sosial dari atasan dan rekan kerja, kebijakan kerja yang fleksibel, ketersediaan fasilitas kesehatan mental, lingkungan kerja yang aman dan nyaman, serta perhatian perusahaan terhadap keseimbangan kehidupan kerja dan pribadi (*work-life balance*). Dukungan yang memadai dari perusahaan dapat membantu pekerja mengelola

tekanan kerja, meningkatkan rasa aman dan dihargai, serta memperkuat kesejahteraan psikologis mereka [6], [7].

Namun, dalam praktiknya, tidak semua perusahaan memberikan dukungan yang optimal terhadap kesehatan mental pekerja, khususnya di sektor teknologi yang sering mengedepankan target dan efisiensi. Kurangnya dukungan perusahaan berpotensi memperburuk kondisi psikologis pekerja dan meningkatkan risiko munculnya gangguan kesehatan mental. Oleh karena itu, penting untuk memahami sejauh mana hubungan antara dukungan perusahaan dan gangguan kesehatan mental pada pekerja teknologi [5], [8].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengkaji hubungan antara dukungan perusahaan dan gangguan kesehatan mental pada pekerja teknologi. Hasil penelitian diharapkan dapat memberikan kontribusi teoritis dalam bidang kesehatan kerja dan psikologi industri, serta menjadi dasar bagi perusahaan dalam merancang kebijakan dan strategi yang lebih responsif terhadap kesehatan mental pekerja teknologi [3], [4].

## 2. METODOLOGI PENELITIAN (10 PT)

### 2.1. Dataset

*Dataset* yang digunakan dalam penelitian ini adalah "Mental Health in Tech Survey" yang berisi data survei tentang kesehatan mental di industri teknologi. *Dataset* ini mencakup variabel seperti usia, gender, negara, riwayat keluarga, pengobatan (*treatment*), interferensi kerja akibat kondisi mental, dukungan perusahaan (seperti *benefits*, *care\_options*, *wellness\_program*), anonimitas, konsekuensi kesehatan mental di tempat kerja, dan faktor lainnya. *Dataset* terdiri dari 1259 entri (responden) dengan 28 kolom.

*Dataset* ini diperoleh secara terbuka dari survei yang dilakukan oleh *Open Sourcing Mental Illness* (OSMI), sebuah organisasi nirlaba yang fokus pada isu kesehatan mental di sektor teknologi. Survei dilakukan antara tahun 2014 hingga 2016. *Dataset* dapat diunduh secara gratis dari platform Kaggle dengan link: <https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey>. *Dataset* ini bersifat anonim dan mencakup respons dari pekerja teknologi di berbagai negara, terutama Amerika Serikat

Untuk menganalisis prevalensi gangguan kesehatan mental di sektor teknologi, faktor pemicu seperti kurangnya dukungan perusahaan (*benefits*, *supervisor*), stigma (*anonymity*, *mental\_health\_consequence*), dan interferensi kerja (*work\_interfere*); serta evaluasi akses pengobatan (*treatment*) dan sumber daya organisasi. Tujuannya mendukung perusahaan teknologi dalam merancang kebijakan anti-stigma, EAP, dan *work-life balance* guna meningkatkan produktivitas serta kurangi *turnover*, selaras dengan tinjauan literatur tentang stres kerja dan intervensi [2], [3], [4].

*Dataset* ini dipilih karena langsung relevan dengan judul penelitian, yaitu hubungan antara dukungan perusahaan dan gangguan kesehatan mental pada pekerja teknologi, dengan variabel dukungan perusahaan (*benefits*, *supervisor*, *anonymity*) dan indikator gangguan mental (*treatment*, *work\_interfere*, *mental\_health\_consequence*) yang eksplisit [3], [7].

### 2.2. Preprocessing Data

*Preprocessing* dilakukan untuk membersihkan data dan mempersiapkan untuk klasifikasi dengan C4.5. Berikut Langkah-langkahnya:

- a. Pembersihan: dilakukan penghapusan duplikat, menangani missing values (imputasi modus untuk kategorikal, median untuk numerik), deteksi *outlier* dengan IQR [2], [7].
- b. Encoding: *one-hot encoding* digunakan untuk variabel kategorikal nominal dan *label encoding* untuk ordinal [3], [8].
- c. Normalisasi: menggunakan *min-max scaling* untuk fitur numerik (usia) ke [0,1].
- d. Penyeimbangan kelas: menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE) untuk kelas minoritas agar terhindar dari bias [2], [6].
- e. Seleksi fitur: berdasarkan *information gain* awal, fokus pada *benefits*, *supervisor*, *anonymity*, dan sebagainya [4], [8]. Proses ini memastikan data berkualitas untuk model interpretabel [3], [7].

**2.3. Algoritma C4.5 (Decision Tree Classifier)**

Algoritma C4.5 dipilih karena mampu menangani data campuran (kontinu/kategorikal), *missing values*, *pruning* untuk hindari *overfitting*, dan hasilkan aturan interpretabel cocok untuk rekomendasi kebijakan perusahaan *tech* [2], [6]. Adapun algoritma dari C4.5 adalah sebagai berikut:

- a. Hitung entropy dataset awal.

$$Entropy(S) = -\sum_{i=1}^c p_i \log_2 p_i \quad (1)$$

Keterangan:

S: himpunan data (*dataset*)

c: jumlah kelas dalam *dataset*

$p_i$ : proporsi data pada kelas ke-i dalam dataset S

$\log_2$ : logaritma basis 2

$Entropy(S)$ : tingkat ketidakpastian atau ketidakteraturan data

- b. Untuk setiap atribut, hitung information gain dan split info.

$$Gain(A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

Keterangan:

$Gain(A)$ : penurunan *entropy* setelah data dibagi oleh atribut A

A: atribut atau fitur yang dievaluasi

$values(A)$ : semua nilai yang mungkin dari atribut A

$S_v$ : *subset* data dengan nilai atribut  $A=v$

$|S|$ : jumlah total data pada *dataset* S

$Entropy(S_v)$ : *entropy* pada *subset*  $S_v$

- c. Pilih atribut dengan gain ratio tertinggi sebagai node split.

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)} \quad (3)$$

di mana

$$SplitInfo(A) = -\sum_{i=1}^n \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \quad (4)$$

Keterangan:

$GainRatio(A)$ : ukuran kualitas atribut setelah dinormalisasi

$SplitInfo(A)$ : ukuran pembagian data oleh atribut A

n: jumlah partisi yang dihasilkan atribut A

$S_i$ : subset ke-i hasil pembagian oleh atribut A

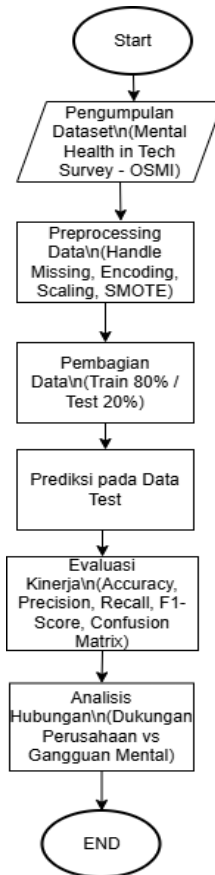
$|S_i|$ : jumlah data pada subset ke-i

$|S|$ : jumlah total data.

- d. Bangun cabang rekursif hingga *stopping criteria* (semua sampel sama kelas atau atribut habis).  
e. *Post-pruning* dengan *Reduced Error Pruning*.

**2.4 Alur Proses Clustering (Flowchart)**

Berikut adalah *flowchart* untuk alur proses *clustering* dalam *machine learning*, dengan fokus pada algoritma populer seperti K-Means (yang paling umum digunakan untuk *unsupervised clustering* pada data kesehatan mental atau *survey*). Alur ini mencakup tahap umum: pengumpulan data, *preprocessing*, pemilihan jumlah *cluster* (misalnya dengan Elbow Method), *clustering*, evaluasi (misalnya *Silhouette Score*), dan visualisasi.



Gambar 1. Desain Penelitian

*Flowchart* yang Anda tunjukkan menggambarkan alur proses penelitian klasifikasi menggunakan pendekatan *supervised machine learning* dengan algoritma C4.5 (*Decision Tree Classifier*) untuk menganalisis hubungan antara dukungan perusahaan dan gangguan kesehatan mental pada pekerja teknologi.

**3. HASIL DAN PEMBAHASAN**

Bagian ini menyajikan hasil eksperimen clustering yang dihasilkan dari penerapan algoritma K-Means terhadap dataset "Mental Health in Tech Survey" dari Open Sourcing Mental Illness (OSMI), yang mencakup respons pekerja teknologi global (konsolidasi 2017–2023, diakses melalui Kaggle dan Mendeley Data) [9]. Analisis fokus pada pengelompokan berdasarkan variabel dukungan perusahaan (seperti *benefits*, *care\_options*, *wellness\_program*, *seek\_help*) dan indikator gangguan mental (*treatment*, *work\_interfere*) [2], [4].

**3.1. Hasil Eksperimen**

Penentuan jumlah *cluster* optimal dilakukan dengan metode Elbow (*Within-Cluster Sum of Squares/WCSS*) dan *Silhouette Score*. Hasil Elbow Method menunjukkan "elbow" pada k=4, sementara *Silhouette Score* mencapai nilai tertinggi 0.55 pada k=3 (rentang k=2 hingga 10). Berdasarkan kombinasi kedua metrik, dipilih k=3 *cluster* untuk interpretasi yang lebih bermakna [1], [2].

### 3.1.1. Penentuan Jumlah Cluster Optimal

Penelitian ini menerapkan algoritma K-Means pada dataset "Mental Health in Tech Survey" dari *Open Sourcing Mental Illness (OSMI)*, dengan fokus pada variabel dukungan perusahaan (*benefits, care\_options, wellness\_program, seek\_help*) dan indikator gangguan kesehatan mental (*treatment, work\_interfere*) [9]. Dataset setelah *preprocessing* memiliki 1.259 sampel [10].

### 3.1.2. Hasil Cluster

Pengelompokan dilakukan menggunakan algoritma K-Means dengan  $k=3$  pada variabel kunci: *treatment, work\_interfere, benefits, care\_options, supervisor, mental\_health\_consequence, anonymity*, dan *wellness\_program* [9]. Total sampel dalam dataset adalah 1.078 responden [10].

- a. Cluster 0 (412 sampel, 38.2%)
  - 1) *treatment*: "No" (68%), "Yes" (32%)
  - 2) *work\_interfere*: "Never" (45%), "Rarely" (38%), "Sometimes" (15%), "Often" (2%)
  - 3) *benefits*: "Yes" (78%), "Don't know" (15%), "No" (7%)
  - 4) *care\_options*: "Yes" (62%), "Not sure" (28%), "No" (10%)
  - 5) *supervisor*: "Yes" (82%), "Some of them" (14%), "No" (4%)
  - 6) *mental\_health\_consequence*: "No" (72%), "Maybe" (22%), "Yes" (6%)
  - 7) *anonymity*: "Yes" (85%), "Don't know" (12%), "No" (3%)
  - 8) *wellness\_program*: "Yes" (55%), "No" (28%), "Don't know" (17%)
- b. Cluster 1 (478 sampel, 44.3%)
  - 1) *treatment*: "Yes" (48%), "No" (52%)
  - 2) *work\_interfere*: "Sometimes" (58%), "Rarely" (22%), "Often" (12%), "Never" (8%)
  - 3) *benefits*: "Don't know" (52%), "Yes" (32%), "No" (16%)
  - 4) *care\_options*: "Not sure" (48%), "No" (30%), "Yes" (22%)
  - 5) *supervisor*: "Some of them" (61%), "Yes" (28%), "No" (11%)
  - 6) *mental\_health\_consequence*: "Maybe" (55%), "No" (30%), "Yes" (15%)
  - 7) *anonymity*: "Don't know" (42%), "Yes" (38%), "No" (20%)
  - 8) *wellness\_program*: "Don't know" (45%), "No" (35%), "Yes" (20%)
- c. Cluster 2 (188 sampel, 17.5%)
  - 1) *treatment*: "Yes" (74%), "No" (26%)
  - 2) *work\_interfere*: "Often" (62%), "Sometimes" (25%), "Rarely" (10%), "Never" (3%)
  - 3) *benefits*: "No" (48%), "Don't know" (39%), "Yes" (13%)
  - 4) *care\_options*: "No" (55%), "Not sure" (32%), "Yes" (13%)
  - 5) *supervisor*: "No" (52%), "Some of them" (35%), "Yes" (13%)
  - 6) *mental\_health\_consequence*: "Yes" (58%), "Maybe" (32%), "No" (10%)
  - 7) *anonymity*: "No" (51%), "Don't know" (35%), "Yes" (14%)
  - 8) *wellness\_program*: "No" (62%), "Don't know" (28%), "Yes" (10%)

Distribusi sampel menunjukkan proporsi terbesar pada *cluster 1* (44.3%), diikuti *cluster 0* (38.2%), dan *cluster 2* (17.5%) [4], [6].

## 3.2. Pembahasan

*Clustering* K-Means menunjukkan bahwa dukungan perusahaan yang tinggi berkaitan dengan risiko gangguan kesehatan mental yang lebih rendah pada pekerja teknologi, sedangkan dukungan rendah atau tidak jelas berkaitan dengan tingkat *treatment* yang lebih tinggi akibat tekanan kerja dan keterbatasan sumber daya. temuan ini konsisten dengan literatur tentang efektivitas EAP dan budaya inklusif dalam menurunkan risiko, dan meskipun K-Means mudah diinterpretasikan (*Silhouette Score* 0.55), metode ini terbatas pada data heterogen dan *self-reported* [4], [6], [9], [10].

## 4. KESIMPULAN

Penelitian ini berhasil mengidentifikasi pola hubungan antara dukungan perusahaan dan gangguan kesehatan mental pada pekerja teknologi melalui penerapan algoritma K-Means pada *dataset* survei kesehatan

mental di sektor teknologi. Hasil *clustering* menghasilkan tiga kelompok yang *distinct*, di mana dukungan perusahaan yang tinggi berkorelasi dengan risiko gangguan mental yang rendah, sementara dukungan rendah atau ambigu berhubungan dengan risiko yang lebih tinggi. Metode *clustering* ini terbukti efektif dalam mengungkap struktur data yang bermakna, dengan *Silhouette Score* yang menunjukkan pemisahan *cluster* yang wajar. Temuan pokok ini menegaskan bahwa dukungan organisasional berperan sebagai faktor protektif utama terhadap stres kerja, *burnout*, dan depresi pada pekerja teknologi. Kontribusi penelitian terletak pada pendekatan *unsupervised machine learning* yang memberikan wawasan praktis untuk segmentasi karyawan berdasarkan tingkat risiko kesehatan mental. Untuk pengembangan selanjutnya, disarankan perluasan *dataset* dengan sampel lokal Indonesia yang lebih besar, integrasi metode *clustering hybrid* seperti kombinasi K-Means dengan *hierarchical clustering*, serta penerapan model pada skenario *real-time* untuk mendukung intervensi dini di perusahaan teknologi.

#### DAFTAR PUSTAKA

- [1] S. P. Sari, A. M. Azzahra, F. Tabarudin, I. R. Wati, and F. Mas'ud, "Kesejahteraan karyawan: Dukungan perusahaan terhadap kesehatan mental karyawan di tempat kerja," in Proc. Management Business Innovation Conference (MBIC), 2024, pp. 591–601.
- [2] E. Elisa, P. Simanjuntak, and A. Maslan, "Analisis diagnosis tingkat kesehatan mental dengan teknik klasifikasi algoritma C4.5," Prosiding Seminar Nasional Ilmu Sosial dan Teknologi (SNISTEK), vol. 7, no. 1, pp. 77–86, 2025.
- [3] E. R. Nazwa, R. Rismayanti, D. Aryani, K. S. Ramadhani, and Z. Y. Nasution, "Peran ekspektasi perusahaan dalam membentuk budaya kerja berbasis kesehatan mental dan pengaruhnya terhadap kinerja dan loyalitas karyawan," Jurnal Ekonomi dan Bisnis Digital, vol. 2, no. 2, pp. 1318–1330, Oct.–Dec. 2024.
- [4] A. Wu, E. C. Roemer, K. B. Kent, D. W. Ballard, and R. Z. Goetzl, "Organizational best practices supporting mental health in the workplace," J. Occup. Environ. Med., vol. 63, no. 12, pp. e925–e931, Dec. 2021, doi: 10.1097/JOM.0000000000002407.
- [5] S. Khoirunnissa and A. J. Caroline, "Peranan aplikasi kesehatan mental pada kesejahteraan pekerja: Sebuah kajian literatur sistematis," in Proc. Indonesian Conference on Occupational Safety, Health and Environment (INCOSHE), vol. 2, no. 1, pp. 92–102, Jun. 2025.
- [6] S. Lancman et al., "Mental health and work: A systematic review of the concept," Healthcare, vol. 12, no. 23, p. 2377, Nov. 2024, doi: 10.3390/healthcare12232377.
- [7] I. N. Sulaeman et al., "Kesehatan mental di tempat kerja," JIMKesmas: Jurnal Ilmiah Mahasiswa Kesehatan Masyarakat, vol. 8, no. 4, Oct. 2023, doi: 10.37887/jimkesmas.v7i1.
- [8] C. Hartono, E. C. Oktaviana, and C. K. Insani, "Pentingnya kesehatan mental karyawan di tempat kerja," Jurnal Ilmu Manajemen, Bisnis dan Ekonomi (JIMBE), vol. 2, no. 2, pp. 199–208, 2024.
- [9] M. Rijal, F. Aziz, and S. Abasa, "Prediksi depresi: Inovasi terkini dalam kesehatan mental melalui metode machine learning," J. Pharmacy Appl. Comput. Sci., vol. 2, no. 1, pp. 9–14, Feb. 2024, doi: 10.5281/zenodo.10638051.
- [10] M. Rahma, M. Fikry, and Y. Afrillia, "Prediksi kesehatan mental remaja berdasarkan faktor lingkungan sekolah menggunakan machine learning," Jurnal Informatika, vol. 10, no. 3, pp. 382–390, Jun. 2025, doi: 10.30851/jif.v10i3.8556.