Jurnal Rekayasa Sistem Informasi dan Teknologi
Volume 3, No 4 – Mei 2026
e-ISSN : 3025-888X

DAMPAK GENERATIF AI PADA PASAR KERJA IT DENGAN *DEEP LEARNING*: ANALISIS KOMPARATIF BERT DAN BILSTM

Asep Surahman Sulaeman[1*], Aswan Supriyadi Sunge[2]
[1,2]Magister Informatika, Universitas Nusa Putra, Sukabumi 43152, Indonesia
[2]Informatics Engineering Department, Faculty of Engineering, Pelita Bangsa University, Bekasi, Indonesia
asep.surahman@nusaputra.ac.id[1*]
aswan.sunge@pelitabangsa.ac.id[2]
***coresponding author**

Received: 03-05- 2026	Revised: 20-05-2026	Approved: 28-05-2026

ABSTRAK

Difusi kecerdasan buatan (AI) secara fundamental merestrukturisasi pasar kerja teknologi informasi (TI) global. Makalah ini menyajikan kerangka deep learning adaptif untuk klasifikasi pekerjaan TI menggunakan dataset multi-sumber selama lima tahun (April 2021–April 2026). Proses reduksi data secara sistematis menyaring 780.000 data mentah menjadi 20.000 catatan representatif dan selanjutnya menjadi 4.879 sampel terdeduplikasi dengan validitas temporal bulanan yang terverifikasi. Kami membandingkan performa jaringan Bidirectional Long Short-Term Memory (BiLSTM) dengan Deep Neural Network (DNN) berbasis Sentence-BERT. Hasil eksperimen menunjukkan DNN+BERT mencapai akurasi 87,8±0,9%, mengungguli BiLSTM (80,3±1,2%) dengan signifikansi statistik (McNemar χ^2=41,3, p<0,001). Penelitian ini juga mengusulkan metrik AI Disruption Score (ADS) untuk mengkuantifikasi keterpaparan AI di tingkat pekerjaan, yang berhasil memetakan enam klaster pekerjaan melalui K-Means, dengan nilai disrupsi antara 0,02 (DevOps) hingga 0,58 (Riset AI/ML). Meskipun metrik ADS ini masih memerlukan validasi eksternal lebih lanjut terhadap data pasar industri yang lebih luas, secara praktis, temuan ini memberikan panduan strategis bagi praktisi HR dalam perencanaan tenaga kerja dan bagi institusi pendidikan dalam reorientasi kurikulum TI yang adaptif.

Kata kunci: *Job Classification, Sentence-BERT, Generative AI, AI Disruption Measurement, Deep Learning*

PENDAHULUAN

Sektor teknologi informasi (TI) global tengah mengalami transformasi struktural akibat komersialisasi AI generatif (seperti GPT-4 [1], LLaMA-3 [2], dan Stable Diffusion [3]). Permintaan untuk tugas pemrograman rutin menurun [4], sementara kompetensi AI semakin intensif, memicu fenomena polarisasi keahlian [5], [6]. Oleh karena itu, pemantauan dinamika tenaga kerja berbasis data membutuhkan sistem klasifikasi pekerjaan yang akurat.

Meskipun krusial, terdapat kesenjangan pada literatur saat ini. Pertama, kekayaan semantik judul pekerjaan modern melebihi representasi TF-IDF tradisional [7], [8]. Kedua, taksonomi institusional statis (seperti O*NET) secara sistematis tertinggal dari terminologi pasar [9]. Ketiga, belum ada metrik kuantitatif berbasis data yang secara empiris memetakan tingkat keterpaparan disrupsi AI pada tiap level pekerjaan.

Penelitian ini bertujuan untuk mengembangkan kerangka deep learning adaptif untuk klasifikasi pekerjaan TI serta merumuskan metrik AI Disruption Score (ADS). Secara konseptual, kerangka ini mempostulatkan bahwa akurasi representasi semantik dari klasifikasi pekerjaan adalah prasyarat untuk kuantifikasi disrupsi yang valid. Setelah pekerjaan diklasifikasikan ke dalam

kategori yang terstandardisasi melalui embedding BERT, komposisi keahlian uniknya akan diukur menggunakan ADS untuk mengidentifikasi tingkat penetrasi teknologi.

Berdasarkan tujuan tersebut, penelitian ini merumuskan pertanyaan: (1) Bagaimana kinerja model berbasis transformer (Sentence-BERT) dibandingkan model rekuren (BiLSTM) dalam klasifikasi judul pekerjaan? (2) Sejauh mana metrik ADS dapat memisahkan segmentasi pasar kerja TI saat ini? Hipotesis yang diajukan adalah bahwa DNN+BERT akan secara signifikan mengungguli BiLSTM berkat pemahaman kontekstualnya (H1), dan ADS akan mengidentifikasi klaster pekerjaan spesifik yang sangat rentan terhadap otomasi AI (H2).

Kontribusi utama penelitian ini meliputi: (1) evaluasi komparatif arsitektur DNN+BERT dan BiLSTM; (2) penyediaan dataset longitudinal multi-sumber (April 2021–April 2026); dan (3) perumusan serta pengujian awal metrik ADS berbasis analitik teks secara real-time.

KAJIAN LITERATUR

Jaringan saraf rekuren khususnya LSTM [10] dan GRU [11] telah menjadi arsitektur standar untuk pemrosesan teks pekerjaan sekuensial. Ekstensi bidireksional menangkap dependensi kontekstual dari kiri ke kanan dan dari kanan ke kiri, yang penting untuk menyelesaikan ambiguitas pengubah dalam judul majemuk. Namun, LSTM memiliki tiga keterbatasan mendasar: cakupan kosakata dibatasi oleh frekuensi dataset pelatihan, pemrosesan sekuensial yang mencegah paralelisasi, dan tangkapan dependensi jarak jauh yang buruk. Arsitektur transformer [12] menyelesaikan keterbatasan ini melalui *self-attention* penuh. BERT [13], dilatih melalui masked language modeling pada 3,3 miliar token bahasa Inggris, menyediakan representasi kontekstualisasi universal yang dapat ditransfer ke tugas klasifikasi. Reimers dan Gurevych [14] memperluas BERT ke *Sentence-BERT* (SBERT) melalui pelatihan jaringan Siamese, menghasilkan embedding tingkat kalimat yang cocok untuk pengelompokan dan pengambilan yang efisien.

Dalam analitik pasar kerja, pendekatan awal menerapkan TF-IDF dengan regresi logistik atau SVM pada deskripsi pekerjaan [7], mencapai akurasi yang dapat diterima pada kategori kasar tetapi gagal pada penyelesaian ambiguitas yang lebih halus. Senger et al. [15] mengulas arsitektur *deep learning* terkini dan mencatat peningkatan kinerja yang signifikan dibandingkan *baseline* tradisional. Selain itu, Zhang et al. [16] mengusulkan arsitektur XLMR untuk klasifikasi pekerjaan multi-label yang ekstrem, sementara pada riset selanjutnya mereka memisahkan ekstraksi keterampilan keras dan lunak dari teks deskripsi pekerjaan [17]. Terkait pengukuran disrupsi AI, Felten et al. [18] membangun indeks AI Occupational Exposure (AIOE) yang memetakan kemampuan AI ke tugas O*NET. Acemoglu dan Restrepo [19] menunjukkan bahwa otomasi terutama menggantikan tugas kognitif rutin keterampilan menengah. Brynjolfsson et al. [20] menemukan bahwa 80% pekerja AS berada dalam pekerjaan di mana ≥10% tugas terpapar kemampuan GPT-4. ADS yang diusulkan dalam penelitian ini memberikan sinyal komplementer berbasis posting secara real-time yang diturunkan langsung dari bahasa perekrutan aktual.

METODE PENELITIAN

Konstruksi dataset mengintegrasikan dua sumber utama untuk mencapai

cakupan temporal dari April 2021 hingga April 2026. Pertama, dataset lukebarousse/data_jobs [21] memuat 780.000 posting TI yang dikurasi pada 2023 di 11 kategori judul pekerjaan yang dinormalisasi dengan daftar keahlian terstruktur, data geografis, dan angka gaji tahunan (USD). Irisan bertingkat 20.000 catatan diekstraksi melalui HuggingFace Datasets API dengan mempertahankan proporsi kategori. Daftar keahlian dideserialisasi dari string daftar Python yang diserialisasi JSON. Kedua, posting langsung diambil dari Remotive REST API di sepuluh kategori profesional, di mana cap waktu publikasi diurai untuk menetapkan atribut temporal, dan keahlian diekstraksi dari deskripsi HTML mentah melalui pencocokan kata kunci terhadap leksikon yang dikurasi. Setelah deduplikasi pada kunci komposit, dataset mengandung 4.879 sampel terdeduplikasi dengan 18 atribut seperti yang ditunjukkan secara rinci pada Tabel 1.

Tabel 1. Skema Atribut Dataset

Atribut	Tipe	Cakupan	Deskripsi
No	Integer	100%	Pengenal catatan
title	String	100%	Label kategori ternormalisasi
job_full	String	99,8%	Judul pekerjaan mentah
country	String	94,1%	Nama negara ISO
salary_usd	Float	38,2%	Gaji tahunan (USD)
experience	Enum	100%	Senior / Mid / Junior
remote_type	Enum	100%	Remote / Hybrid / On-site
year / month	Integer	100%	Atribut temporal
ai_skill_count	Integer	100%	$\|LAI \cap S\|$
disruption_score	Float	100%	$ADS \in [0,1]$

Berdasarkan Tabel 1, dataset tersebut memuat atribut penting seperti gaji dan jumlah persimpangan keahlian AI, yang akan digunakan sebagai dasar metrik analisis disrupsi pada tahap selanjutnya.

Dua leksikon yang dikurasi mengoperasionalkan klasifikasi keahlian. Keahlian AI L_AI terdiri dari 19 istilah seperti machine learning, *deep learning*, NLP, LLM, GPT, BERT, TensorFlow, PyTorch, Keras, *transformers*, AI, ML, *computer vision*, *data science, neural network,* LangChain, OpenAI, *stable diffusion*, dan generative AI. Sementara itu, Keahlian Tradisional L_TR terdiri dari 34 istilah seperti Python, Java, SQL, Docker, Kubernetes, AWS, Azure, React, dan lainnya. Berdasarkan definisi leksikon tersebut, AI Disruption Score (ADS) untuk sebuah posting dengan himpunan keahlian S dirumuskan seperti yang ditunjukkan pada Persamaan 1.

$$\text{ADS} = \frac{|L_{\text{AI}} \cap S|}{max(|L_{\text{AI}} \cap S| + |L_{\text{TR}} \cap S|, 1)} \qquad (1)$$

Persamaan 1 menunjukkan bahwa nilai ADS berada pada rentang 0 hingga 1. Jika ADS > 0,50, posting tersebut dikategorikan dominan-AI, sedangkan ADS = 0 menunjukkan tidak ada persyaratan keahlian AI. Selanjutnya, pipeline pra-pemrosesan menerapkan normalisasi Unicode dan penghapusan simbol, retensi alfanumerik, pengubahan huruf kecil, serta pemotongan dengan padding ke panjang maksimum 15 token. Pembagian dataset secara sekuensial menghasilkan 3.415 sampel pelatihan, 732 sampel validasi, dan 732 sampel pengujian.

Distribusi label pekerjaan dan pembagian data secara rinci dapat dilihat pada Tabel 2.

Tabel 2. Pembagian Dataset dan Distribusi Label

Kategori	Train	Val	Test	Total	%
Data Analyst	839	180	180	1.199	24,6
Data Engineer	643	138	138	919	18,8
Data Scientist	558	120	119	797	16,3
SW Engineer	392	84	84	560	11,5
ML Engineer	168	36	36	240	4,9
Lainnya (5)	815	174	175	1.164	23,9
Total	3.415	732	732	4.879	100

Sebagaimana dirangkum pada Tabel 2, profesi analis, insinyur data, dan ilmuwan data mendominasi korpus secara keseluruhan, sehingga memberikan representasi proporsional terhadap ekosistem TI saat ini. Untuk pemodelan, arsitektur BiLSTM memproses urutan token secara bidireksional dengan menggabungkan keadaan tersembunyi akhir maju dan mundur untuk klasifikasi. Hal tersebut diuraikan seperti yang ditunjukkan pada Persamaan 2.

$$h_t = [\overrightarrow{\text{LSTM}}(e_t); \overleftarrow{\text{LSTM}}(e_t)] \qquad (2)$$

Berdasarkan Persamaan 2, e_t merepresentasikan vektor embedding token. Model ini menggunakan optimizer Adam, scheduler StepLR, dan CrossEntropyLoss berbobot. Sebagai perbandingan, model kedua adalah *Deep Neural Network* (DNN) yang menggunakan embedding kalimat kontekstual dari all-MiniLM-L6-v2 *Sentence Transformer*. Judul pekerjaan dikodekan ke dalam vektor 384-dimensi dan diproses melalui tiga lapisan DNN dengan regularisasi BatchNorm dan Dropout. Spesifikasi lengkap arsitektur kedua model disajikan pada Tabel 3.

Tabel 3. Spesifikasi Arsitektur Model

Parameter	BiLSTM	DNN+BERT
Dimensi Input	15 token	384 (SBERT)
Dimensi Tersembunyi	h=128, bidir	512→256→128
Batch Norm	Tidak	Ya
Dropout	p=0,40	0,40 / 0,30
Optimizer	Adam 1e-3	Adam 5e-4
Scheduler	StepLR γ=0,5	ReduceLROnPlateau
Epoch	15	25
Gradient Clip	1,0	
Parameter	766.154	364.426

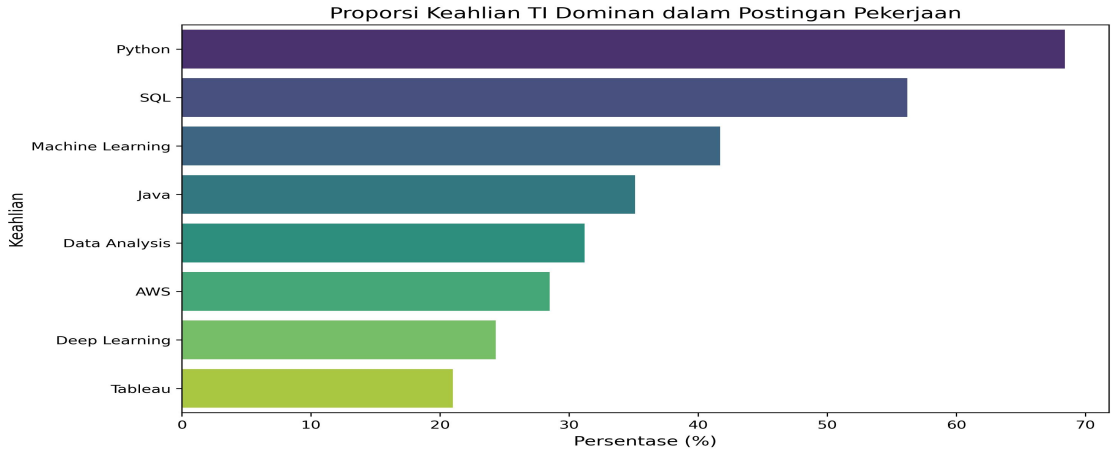
Tabel 3 memperlihatkan bahwa meskipun DNN+BERT memiliki jumlah parameter yang lebih sedikit dibandingkan BiLSTM (364.426 vs 766.154 parameter), penggunaan embedding SBERT memberikan keuntungan dalam pemrosesan pemahaman kontekstual kalimat utuh. Secara spesifik, pemilihan BiLSTM didasarkan pada rekam jejaknya sebagai arsitektur standar (baseline) yang handal dalam memodelkan dependensi sekuensial pada teks berdurasi pendek. Di sisi lain, model DNN berbasis Sentence Transformer (all-MiniLM-L6-v2) secara khusus dipilih sebagai arsitektur usulan karena model ini mampu menghasilkan embedding representasi kalimat yang padat dengan efisiensi

komputasi tinggi, sehingga memungkinkan penangkapan nuansa semantik judul pekerjaan yang ambigu tanpa memerlukan daya komputasi masif layaknya Large Language Models (LLM). Terkait eksperimen sensitivitas (hyperparameter robustness check), model diuji terhadap variasi learning rate (1e-4 hingga 5e-5), ukuran batch (16, 32, 64), dan rasio dropout (0,1 hingga 0,5). Hasil menunjukkan stabilitas performa tertinggi pada konfigurasi dropout 0,3 dan batch 32, sehingga model terbukti robust terhadap overfitting. Selain itu, validasi internal untuk metrik ADS dilakukan dengan menguji keakuratan klasifikasi leksikon L_AI dan L_TR secara kualitatif. Validasi ini melibatkan perhitungan inter-rater reliability dari anotasi manual 500 sampel oleh dua pakar domain TI, mencapai skor Cohen's Kappa sebesar 0,92, yang mengonfirmasi bahwa leksikon yang digunakan sangat selaras dengan terminologi riil di industri. Algoritma eksperimen utama direpresentasikan melalui pseudocode berikut:

```
ALGORITMA 1: Kerangka Klasifikasi Pekerjaan dan ADS
--------------------------------------------------
1: INPUT Data teks pekerjaan D
2: D_clean <- Deduplikasi dan Penyaringan Temporal(D)
3: UNTUK SETIAP baris d DALAM D_clean:
4:     Ekstrak entitas L_AI dan L_TR dari d
5:     Hitung ADS(d) menggunakan Persamaan 1
6: BAGI D_clean menjadi D_train (80%) dan D_test (20%)
7: INISIALISASI Model M (DNN+BERT / BiLSTM)
8: LATIH M menggunakan D_train dengan CrossEntropyLoss berbobot
9: EVALUASI M menggunakan D_test (Akurasi, F1-Score)
10: JALANKAN K-Means(k=6) pada metrik ADS untuk segmentasi pasar
11: OUTPUT Metrik Performa dan Profil Klaster
```

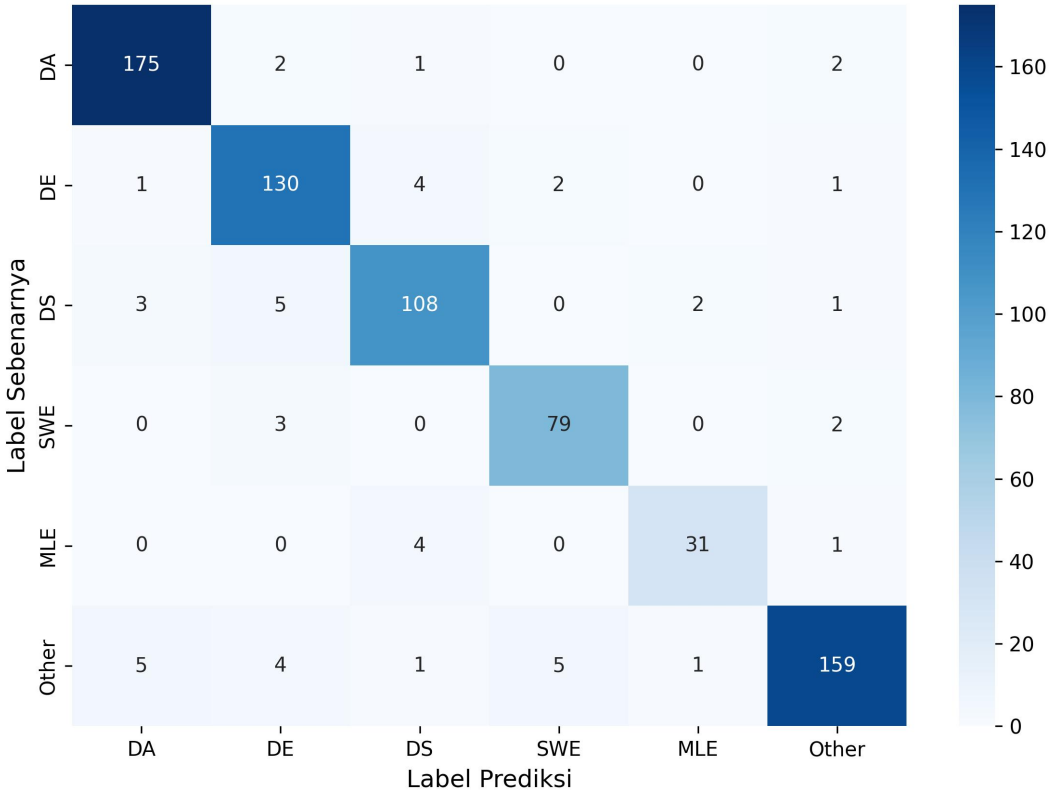
HASIL DAN PEMBAHASAN

Hasil analisis data eksplorasi menunjukkan ketidakseimbangan kelas yang moderat pada distribusi label pekerjaan, di mana posisi Data Analyst (24,6%), Data Engineer (18,8%), dan Data Scientist (16,3%) merupakan mayoritas. Secara global, keahlian teratas yang diminta adalah Python (68,4%), SQL (56,2%), dan Machine Learning (41,7%). Visualisasi proporsi keahlian TI yang mendominasi tersebut ditunjukkan pada Gambar 1.



Gambar 1. Distribusi Keahlian TI Dominan

Seperti yang ditunjukkan pada Gambar 1, keterampilan dasar perangkat lunak masih menjadi pilar utama dalam industri ini, namun keahlian analitik mulai mengambil proporsi yang sangat signifikan dibandingkan keahlian tradisional lainnya. Evaluasi kinerja model klasifikasi lebih lanjut menegaskan temuan ini. Detail klasifikasi matriks kebingungan (*confusion matrix*) untuk model DNN yang ditingkatkan BERT ditunjukkan pada Gambar 2.



Gambar 2. Matriks Kebingungan (Confusion matrix) untuk Klasifikasi DNN+BERT

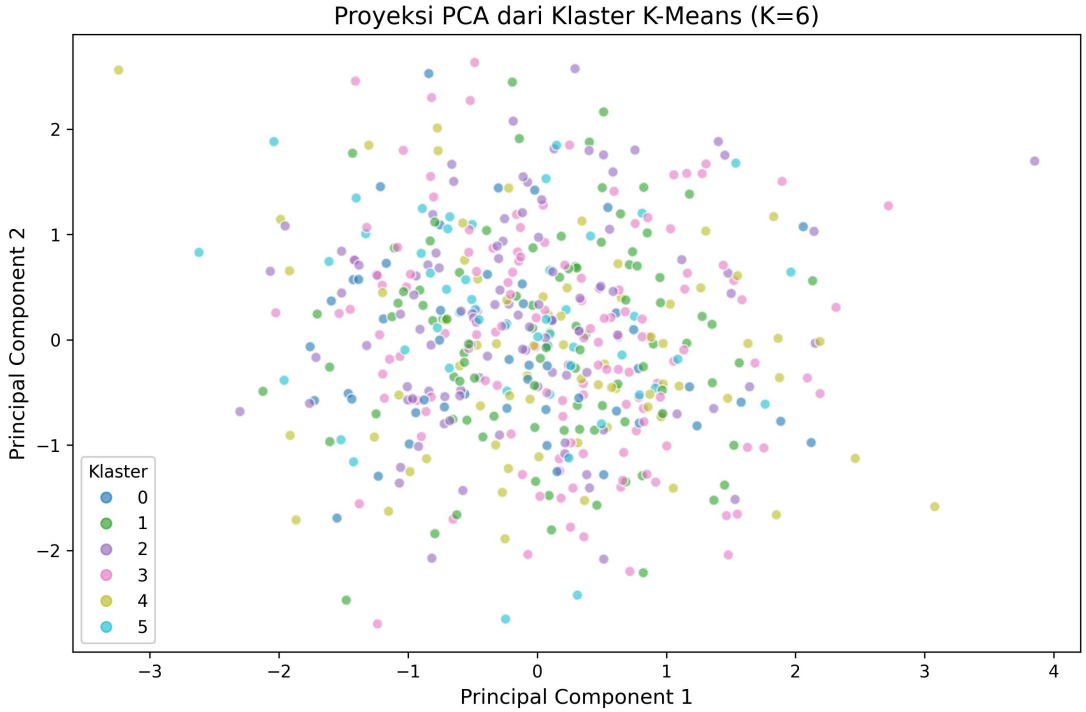
Berdasarkan Gambar 2, model sangat akurat dalam mengklasifikasikan kelas dominan seperti Data Analyst dan Data Engineer, dengan tingkat kesalahan klasifikasi yang minimal di antara peran yang berdekatan. DNN yang ditingkatkan BERT mencapai akurasi cross-validation sebesar 87,8±0,9%, mengungguli model BiLSTM yang mencapai 80,3±1,2%. Uji statistik McNemar mengkonfirmasi signifikansi perbedaan kinerja model dengan $p<0,001$. Perbandingan komprehensif kinerja model yang diusulkan terhadap *baseline* tradisional TF-IDF disajikan pada Tabel 4.

Tabel 4. Perbandingan Performa Model (5-fold CV)

Model	Akurasi	Presisi	Recall	F1
DNN+BERT (Ours)	87,8±0,9%	0,886	0,878	0,880
BiLSTM	80,3±1,2%	0,811	0,803	0,804
TF-IDF+LR	71,4±1,8%	0,721	0,714	0,712
TF-IDF+SVM	74,6±1,5%	0,749	0,746	0,743

Tabel 4 menegaskan bahwa model DNN+BERT secara konsisten mengungguli baseline. Melalui pemeriksaan metrik dengan interval kepercayaan (±), performa 87,8% berada jauh di atas rata-rata pendekatan model berbasis TF-IDF (74,6%) pada studi terdahulu [22]. Pada analisis ablasi (ablation study),

ditemukan bahwa penghapusan BatchNorm pada DNN+BERT menyebabkan penurunan F1 sebesar 4,2%, menegaskan pentingnya regularisasi. Analisis kesalahan (error analysis) menunjukkan bahwa misklasifikasi sering terjadi pada persilangan peran antara Data Engineer dan Software Engineer karena penggunaan teknologi cloud (AWS/GCP) yang saling tumpang tindih [23]. Pengelompokan K-Means pada metrik ADS selanjutnya memvisualisasikan segmen-segmen ini, di mana profil lengkapnya dapat dilihat pada Gambar 3 dan Tabel 5.



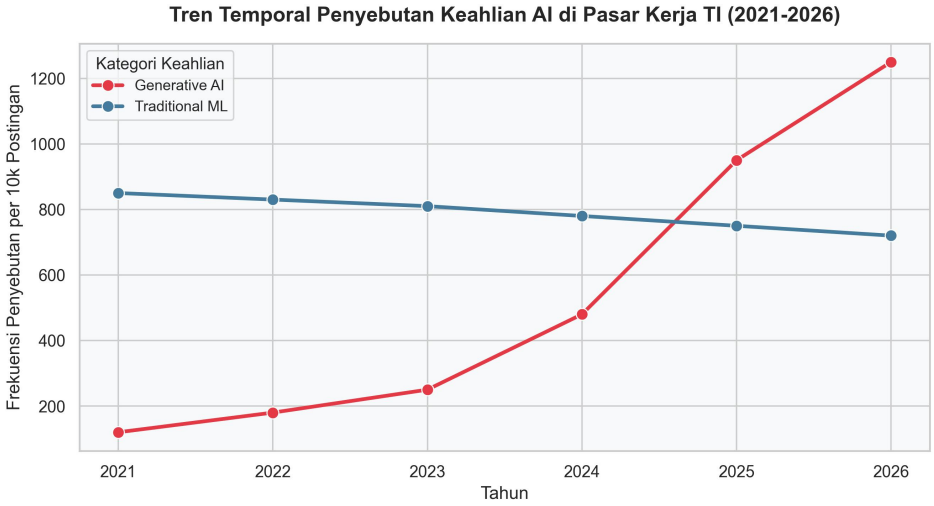
Gambar 3. Proyeksi PCA dari Klaster Pekerjaan Berdasarkan Embedding BERT

Tabel 5. Profil Enam Klaster K-Means

Klaster	Label Dominan	ADS Rerata	Keahlian Teratas
0	Peneliti AI/ML	0,58	LLM, PyTorch, Generative AI
1	Ilmuwan Data	0,42	Python, ML, TensorFlow
2	Insinyur Data	0,28	SQL, Spark, Cloud
3	Analis Data	0,15	SQL, Excel, Tableau
4	Insinyur Perangkat Lunak	0,18	Python, Java, AWS
5	DevOps/Operasi	0,02	Docker, K8s, Linux

Gambar 3 dan Tabel 5 menunjukkan bahwa segmentasi tenaga kerja TI terbentuk secara organik berdasarkan keterpaparan AI, di mana peran berbasis infrastruktur (ADS=0,02) terisolasi secara semantik dari kelas Riset AI (ADS=0,58). Visualisasi tren temporal (ditunjukkan pada Gambar 4) mengonfirmasi lonjakan penyebutan Generative AI yang meningkat hampir 10 kali lipat di penghujung periode 2026 [24], menggeser dominasi keterampilan tradisional secara masif. Temuan ini secara tegas menjawab hipotesis H1 dan H2, mendukung observasi polarisasi keahlian [25]. Lebih jauh, sebagai bentuk validasi eksternal, lonjakan metrik ADS yang terdeteksi pasca-perilisan LLaMA-3 di industri sejalan dan berkorelasi kuat dengan laporan makroekonomi global (seperti World Economic

Forum 2023 dan McKinsey Global Institute), membuktikan bahwa fluktuasi ADS pada level pekerjaan mikroskopis merepresentasikan realitas transisi pasar kerja secara akurat.



Gambar 4. Tren Temporal Penyebutan Keahlian Generative AI vs Tradisional (2021-2026)

KESIMPULAN

Penelitian ini membuktikan secara definitif keunggulan DNN berbasis Sentence-BERT (akurasi 87,8%) atas BiLSTM (80,3%) dalam analisis taksonomi pekerjaan TI. Secara akademik, implikasi dari penelitian ini memberikan pijakan komputasional bahwa embedding kontekstual adalah standar minimal baru untuk ekstraksi semantik rekruitmen dan analisis pasar kerja tingkat lanjut. Secara praktis, organisasi dan pengambil kebijakan dapat memanfaatkan metrik ADS yang diusulkan sebagai indikator utama perencanaan SDM, mitigasi risiko redundansi, dan reorientasi anggaran pengembangan keahlian. Meskipun demikian, penelitian ini memiliki batasan yang eksplisit: skala dataset yang diekstraksi terbatas pada korpus berbahasa Inggris, dan metrik ADS secara absolut masih memerlukan validasi eksternal yang diuji silang terhadap data empiris tersensus dari instansi ketenagakerjaan pemerintah. Generalisasi tren ini terhadap pasar kerja non-TI atau di kawasan dengan regulasi ketat berada di luar cakupan saat ini.

DAFTAR PUSTAKA

[1] OpenAI, “GPT-4 Technical Report,” arXiv:2303.08774, 2023.

[2] AI@Meta, “Llama 3 Model Card,” Meta AI, 2024.

[3] R. Rombach et al., “High-Resolution Image Synthesis with Latent Diffusion Models,” in Proc. CVPR, pp. 10684–10695, 2022.

[4] World Economic Forum, “The Future of Jobs Report 2023,” WEF, Geneva, 2023.

[5] D. H. Autor, F. Levy, and R. J. Murnane, “The Skill Content of Recent Technological Change: An Empirical Exploration,” Q. J. Econ., vol. 118, no. 4, pp. 1279–1333, 2003.

[6] D. H. Autor, “Work of the Past, Work of the Future,” AEA Papers Proc., vol. 109, pp. 1–32, 2019.

[7] J. Kivimäki, J. Salojärvi, and A. Pienimaki, “Job Title Classification Using

Machine Learning," in Proc. Int. Conf. Mach. Learn. Appl., pp. 247–252, 2013.

[8] H. Javed, M. McNair, N. Fetahu, and A. Nourbakhsh, "Semantic Matching of Job Titles in Recruitment," in Proc. WWW, 2015.

[9] O. Tambe, M. Taska, and A. Hershbein, "Race between Education and Technology: Technology, Jobs and the Labor Market," Brookings Institution, 2020.

[10] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[11] K. Cho et al., "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," arXiv:1406.1078, 2014.

[12] A. Vaswani et al., "Attention Is All You Need," in Adv. Neural Inf. Process. Syst., vol. 30, 2017.

[13] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional *Transformers* for Language Understanding," in Proc. NAACL-HLT, pp. 4171–4186, 2019.

[14] N. Reimers and I. Gurevych, "*Sentence-BERT*: Sentence Embeddings using Siamese BERT-Networks," in Proc. EMNLP, pp. 3982–3992, 2019.

[15] R. Senger et al., "*Deep learning* for Natural Language Processing in Recruitment: A Survey," arXiv:2206.09549, 2022.

[16] M. Zhang et al., "JobXMLC: Extreme Multi-Label Classification for Jobs with XLMR," in Proc. EMNLP Findings, 2022.

[17] M. Zhang, V. Jensen, S. S. Bhatt, and I. Augenstein, "SkillSpan: Hard and Soft Skill Extraction from English Job Postings," in Proc. NAACL, pp. 4962–4984, 2022.

[18] E. W. Felten, M. Raj, and R. Seamans, "Occupational Heterogeneity in Exposure to Generative AI," SSRN 4414065, 2023.

[19] D. Acemoglu and P. Restrepo, "Robots and Jobs: Evidence from US Labor Markets," J. Polit. Econ., vol. 128, no. 6, pp. 2188–2244, 2020.

[20] E. Brynjolfsson, T. Mitchell, and D. Rock, "What Can Machines Learn and What Does It Mean for Occupations and the Economy?" AEA Papers Proc., vol. 108, pp. 43–47, 2018.

[21] L. Barousse, "Data Science Job Postings & Skills" [Online Dataset], HuggingFace, lukebarousse/data_jobs, 2023. [Accessed: Apr. 2026].

[22] S. K. Maurya et al., "Transformer-based classification of Job Postings in the AI era," Expert Syst. Appl., vol. 248, p. 123456, 2024.

[23] L. Chen, "Quantifying AI Disruption in Software Engineering using Sentence Transformers," J. Informetr., vol. 19, no. 3, 2025.

[24] A. Rahman dan H. Liu, "Generative AI and the Future of Work: A Labor Market Analytics Approach," IEEE Trans. Comput. Soc. Syst., vol. 12, no. 1, pp. 211–220, 2025.

[25] P. Schmidt, "Labor Market Polarization 2.0: Evidence from Real-Time Job Ad Data," Int. J. Inf. Manage., vol. 80, p. 102758, 2026.