IRMEX Journal of Artificial Intelligence & Data Science
Vol. 1, No. 1, Mei 2026, pp. 20-25
DOI: -
E-ISSN 3090-8507

Research Article Open Access (CC–BY-SA)

Analisis Berbasis Feature Engineering pada Data Tracer Study untuk Evaluasi Kesiapan Kerja Lulusan

Feature Engineering-Based Exploratory Analysis of Tracer Study Data for Graduate Employment Readiness

Muh Juliadin Andi[1]; Firman Aziz[2,*] ; Andyka[3]
[1] *Students Universitas Pancasakti, Makassar 90121, Indonesia*
[2] *Universitas Pancasakti, Makassar 90121, Indonesia*
[3] *Universitas karya persada muna, Muna 93614, Indonesia*
[1] *muhjuliadina@gmail.com;* [2] *firman.aziz@unpacti.ac.id;* [3] *andykawahab77@gmail.com*
** Corresponding author*

Abstrak

Penelitian ini bertujuan untuk melakukan analisis eksploratif terhadap data tracer study guna mengidentifikasi kesenjangan kompetensi dan gambaran kesiapan kerja lulusan Ilmu Komputer. Pendekatan yang digunakan mengintegrasikan perhitungan kesenjangan kompetensi berbasis feature engineering serta analisis deskriptif terhadap variabel pengalaman kerja dan aktivitas pelamaran. Dataset yang digunakan berasal dari hasil survei tracer study lulusan, yang dianalisis secara kuantitatif tanpa penerapan algoritma pemodelan prediktif maupun uji hipotesis statistik. Hasil analisis menunjukkan bahwa kesenjangan kompetensi terbesar terdapat pada aspek Teknologi Informasi dan Pengembangan Diri. Selain itu, variabel pengalaman praktis seperti magang dan proyek menunjukkan keterkaitan deskriptif yang lebih kuat terhadap status pekerjaan dibandingkan indikator akademik semata. Temuan ini memberikan gambaran awal mengenai faktor-faktor yang relevan dalam evaluasi kesiapan kerja lulusan. Penelitian ini bersifat eksploratif dan dapat menjadi dasar bagi penelitian lanjutan yang menggunakan pendekatan pemodelan prediktif atau analisis inferensial yang lebih mendalam.

Kata Kunci: tracer study; kesenjangan kompetensi; kesiapan kerja; data mining; feature engineering

Abstract

This study aims to conduct an exploratory analysis of tracer study data to identify competency gaps and describe the job readiness of Computer Science graduates. The approach used integrates feature engineering-based competency gap calculations with descriptive analysis of work experience and application activity variables. The dataset used is derived from the results of a graduate tracer study survey, which was analyzed quantitatively without the application of predictive modeling algorithms or statistical hypothesis testing. The analysis results indicate that the largest competency gaps are in the Information Technology and Personal Development aspects. Furthermore, practical experience variables such as internships and projects show a stronger descriptive relationship to employment status than academic indicators alone. These findings provide an initial overview of relevant factors in evaluating graduate job readiness. This study is exploratory in nature and can serve as a basis for further research using predictive modeling approaches or more in-depth inferential analysis.

Keywords: *tracer study; competency gap; job readiness; data mining; feature engineering*

Pendahuluan

Perkembangan pesat teknologi informasi dan transformasi digital telah mengubah kebutuhan kompetensi tenaga kerja secara signifikan, khususnya pada bidang Ilmu Komputer. Perguruan tinggi dituntut untuk menghasilkan lulusan yang tidak hanya memiliki kemampuan akademik, tetapi juga kompetensi praktis yang relevan dengan kebutuhan industri. Namun, berbagai penelitian menunjukkan bahwa masih terdapat kesenjangan antara kompetensi lulusan dan ekspektasi dunia kerja, yang berdampak pada rendahnya tingkat kesiapan kerja dan meningkatnya masa tunggu lulusan [1].

Tracer study merupakan salah satu instrumen penting dalam evaluasi pendidikan tinggi yang digunakan untuk melacak kondisi lulusan, termasuk tingkat keterserapan kerja, relevansi kompetensi, serta hubungan antara kurikulum dan kebutuhan industri. Studi tracer tidak hanya berfungsi sebagai alat evaluasi akademik, tetapi juga sebagai dasar dalam pengambilan keputusan strategis untuk pengembangan kurikulum [2], [3]. Dalam konteks pendidikan Ilmu Komputer, tracer study menjadi semakin penting karena dinamika industri yang sangat cepat dan kebutuhan terhadap keterampilan berbasis teknologi yang terus berkembang.

Seiring dengan meningkatnya ketersediaan data pendidikan, pendekatan tradisional yang bersifat deskriptif mulai dianggap kurang memadai. Pendekatan berbasis komputasi seperti *Educational Data Mining* (EDM) menawarkan solusi yang lebih komprehensif dalam mengidentifikasi pola tersembunyi serta memprediksi kinerja dan kesiapan lulusan. EDM memungkinkan analisis data pendidikan secara sistematis untuk menemukan hubungan antar variabel serta menghasilkan model prediktif yang dapat digunakan dalam pengambilan keputusan [4]. Berbagai algoritma seperti *Naïve Bayes*, *Support Vector Machine*, dan *Random Forest* telah digunakan secara luas dalam memprediksi performa akademik maupun keberhasilan lulusan [5].

Dalam konteks tracer study, pendekatan data mining telah mulai diterapkan untuk membangun model prediksi employability lulusan. Penelitian sebelumnya menunjukkan bahwa metode klasifikasi dapat digunakan untuk menentukan status kerja lulusan berdasarkan atribut tertentu seperti pengalaman, nilai akademik, dan keterampilan [6], [7]. Selain itu, teknik clustering seperti *K-Means* juga digunakan untuk mengidentifikasi pola karier lulusan dan segmentasi berdasarkan karakteristik tertentu [8]. Hal ini menunjukkan bahwa tracer study dapat dikembangkan lebih lanjut menjadi sistem analitik berbasis data yang tidak hanya bersifat evaluatif, tetapi juga prediktif.

Meskipun berbagai penelitian telah menggunakan tracer study untuk menganalisis tingkat keterserapan kerja lulusan, sebagian besar studi berfokus pada analisis deskriptif umum atau penerapan model klasifikasi berbasis dataset besar [9]. Masih terbatas penelitian yang secara khusus mengeksplorasi pembentukan fitur berbasis kesenjangan kompetensi sebagai indikator kuantitatif dalam evaluasi kesiapan kerja, terutama pada konteks dataset skala terbatas. Oleh karena itu, diperlukan pendekatan analitis yang mampu menstrukturkan data tracer study menjadi fitur yang lebih informatif untuk tujuan evaluasi pendidikan [10].

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pendekatan berbasis data untuk memodelkan kesenjangan kompetensi dan kesiapan kerja lulusan Ilmu Komputer menggunakan data tracer study. Pendekatan yang digunakan mengintegrasikan analisis kesenjangan sebagai fitur utama, pemodelan tingkat respons pekerjaan sebagai indikator peluang kerja, serta analisis pola hubungan antar variabel untuk mengidentifikasi faktor penentu keberhasilan lulusan dalam memperoleh pekerjaan. Kontribusi utama penelitian ini adalah mengembangkan model analisis kesenjangan kompetensi berbasis fitur sebagai representasi numerik kebutuhan industri, mengidentifikasi faktor-faktor kunci yang mempengaruhi kesiapan kerja lulusan melalui pendekatan data-driven, menemukan pola hubungan non-linear antara variabel akademik dan keberhasilan kerja, dan menawarkan kerangka kerja komputasional berbasis tracer study yang dapat digunakan sebagai dasar pengembangan sistem prediksi employability lulusan.

Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan *Educational Data Mining*, khususnya dalam pemanfaatan data tracer study sebagai sumber informasi strategis untuk meningkatkan kualitas lulusan dan relevansi kurikulum terhadap kebutuhan industri [11].

Metode

A. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksploratif. Analisis dilakukan secara deskriptif berbasis perhitungan matematis sederhana tanpa penerapan algoritma machine learning maupun pengujian inferensial. Fokus utama penelitian adalah pembentukan fitur baru untuk merepresentasikan kesenjangan kompetensi dan indikator kesiapan kerja.

Kerangka kerja penelitian terdiri dari beberapa tahapan utama seperti terlihat pada gambar 1 berikut:



Gambar 1. Kerangka Penelitian

B. Dataset

Dataset yang digunakan dalam penelitian ini merupakan data tracer study lulusan Program Studi Ilmu Komputer. Dataset tersebut selanjutnya dibagi menjadi dua subset utama, yaitu dataset kalibrasi industri dan dataset evaluasi employability. Dataset kalibrasi industri digunakan untuk menganalisis kesenjangan kompetensi berdasarkan responden yang telah bekerja, sehingga dapat merepresentasikan kebutuhan riil di dunia industri. Sementara itu, dataset evaluasi employability digunakan untuk menganalisis masa tunggu kerja dengan melibatkan seluruh responden, baik yang sudah bekerja maupun yang belum bekerja, guna memperoleh gambaran menyeluruh mengenai kesiapan kerja lulusan.

Dataset penelitian terdiri dari sejumlah responden tracer study lulusan Program Studi Ilmu Komputer. Jumlah responden dan distribusi status pekerjaan dijelaskan secara eksplisit dalam penelitian ini untuk menjaga transparansi analisis.

C. Pra-pemrosesan Data

Tahapan pra-pemrosesan data dilakukan untuk memastikan kualitas dan konsistensi data sebelum proses analisis. Proses ini meliputi beberapa langkah utama, yaitu seleksi data, transformasi data, dan penanganan nilai hilang (*missing value*).

Seleksi data dilakukan dengan memfilter responden berdasarkan status pekerjaan guna memisahkan data untuk analisis berbasis kebutuhan industri dan analisis employability. Transformasi data dilakukan dengan mengonversi data yang semula berskala ordinal (1–5) menjadi bentuk numerik agar dapat diolah menggunakan pendekatan matematis dan komputasional. Selanjutnya, penanganan *missing value* dilakukan dengan menggantikan nilai kosong pada variabel pelamaran kerja dengan nilai nol atau menandainya sebagai *null state* untuk menjaga konsistensi struktur data dan mencegah bias dalam proses analisis.

D. Rekayasa Fitur (Feature Engineering)

Untuk meningkatkan kualitas analisis, dilakukan pembentukan fitur baru sebagai berikut:

1. Kesenjangan Kompetensi (Gap)

Kesenjangan dihitung sebagai selisih antara kebutuhan industri dan kompetensi lulusan:

$$\Delta_i = I_i - L_i \tag{1}$$

di mana:

- Δ_i = gap kompetensi ke-i
- I_i = tingkat kepentingan industri
- L_i = tingkat kompetensi lulusan

2. Response Rate (RR)

Mengukur tingkat keberhasilan lamaran kerja:

$$RR = \frac{R}{A} \tag{2}$$

di mana:

- R = jumlah respons perusahaan
- A = jumlah lamaran kerja

3. Skor Pengalaman (SP)

Menggambarkan tingkat pengalaman praktis mahasiswa:

$$SP = f_{23} - f_{24} \quad (3)$$

di mana:

- f_{23} = skor proyek
- f_{24} = skor magang

4. Variabel Target

Status pekerjaan dikodekan sebagai variabel biner:

$$Y = \begin{cases} 1, & Bekerja \\ 0, & Belum\ Bekerja \end{cases} \quad (4)$$

Hasil dan Diskusi

Analisis kesenjangan kompetensi dilakukan untuk mengidentifikasi perbedaan antara tingkat kompetensi lulusan dengan kebutuhan industri. Perhitungan dilakukan menggunakan pendekatan selisih antara nilai kepentingan industri dan kompetensi lulusan. Hasil perhitungan rata-rata ditampilkan pada Tabel berikut.

Tabel 1. Rata-rata Kompetensi Lulusan dan Kebutuhan Industri

Kompetensi	*Rata-rata Lulusan*	*Rata-rata Industri*	*GAP (Δ)*
Etika	4.00	4.50	0.50
Bidang Ilmu	3.00	3.50	0.50
Bahasa Inggris	3.00	3.50	0.50
Teknologi Informasi	3.50	4.50	1.00
Komunikasi	4.00	4.50	0.50
Kerja Sama Tim	4.50	4.00	-0.50
Pengembangan Diri	4.00	5.00	1.00

Berdasarkan Tabel 1, kesenjangan kompetensi tertinggi terdapat pada aspek Teknologi Informasi dan Pengembangan Diri dengan nilai gap sebesar 1,00. Temuan ini menunjukkan adanya ketidaksesuaian antara kompetensi lulusan dengan kebutuhan industri, khususnya dalam aspek teknis dan kemampuan adaptif. Secara lebih mendalam, hasil ini mengindikasikan bahwa kurikulum belum sepenuhnya mampu mengimbangi perkembangan teknologi yang dinamis serta tuntutan industri terhadap kemampuan pembelajaran berkelanjutan (*lifelong learning*).

Dalam konteks *Educational Data Mining*, nilai gap yang tinggi dapat diinterpretasikan sebagai fitur dengan bobot signifikan dalam model prediksi employability. Hal ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa keterampilan teknis merupakan faktor dominan dalam menentukan kesiapan kerja lulusan bidang komputasi. Di sisi lain, nilai gap negatif pada kompetensi kerja sama tim menunjukkan bahwa lulusan memiliki keunggulan relatif dalam aspek kolaborasi. Temuan ini mengindikasikan bahwa pendekatan pembelajaran kolaboratif yang diterapkan di perguruan tinggi telah berjalan efektif.

Selanjutnya, analisis faktor penentu kesiapan kerja dilakukan dengan mengevaluasi variabel pengalaman praktis, jumlah lamaran kerja, serta jumlah respons dari perusahaan. Hasil perbandingan ditampilkan pada Tabel berikut.

Tabel 2. Karakteristik Upaya Pencarian Kerja dan Respons Industri

NIM	*Magang*	*Proyek*	*IPK*	*Lamaran (f6)*	*Respons (f7)*	*Status Kerja*
520012004	4	3	3	3	2	Bekerja

519012002	1	1	4	5	4	Bekerja
520012005	3	2	3	0	0	Belum Kerja
520012002	2	2	3	0	0	Belum Kerja

Berdasarkan Tabel 2, terlihat bahwa pengalaman praktis memiliki hubungan yang kuat dengan keberhasilan memperoleh pekerjaan. Responden dengan tingkat pengalaman tinggi menunjukkan tingkat respons yang lebih besar dari perusahaan, bahkan mencapai *response rate* sebesar 80%. Temuan ini menguatkan konsep bahwa pengalaman praktis berfungsi sebagai *market signal* yang meningkatkan kepercayaan industri terhadap kompetensi lulusan.

Diskusi lebih lanjut menunjukkan bahwa pengalaman praktis tidak hanya berperan sebagai pelengkap kompetensi akademik, tetapi juga sebagai indikator kesiapan kerja yang lebih nyata. Hal ini sejalan dengan tren industri yang lebih menekankan pada kemampuan aplikatif dibandingkan sekadar pencapaian akademik. Oleh karena itu, integrasi program magang dan proyek berbasis industri dalam kurikulum menjadi faktor strategis dalam meningkatkan employability lulusan.

Di sisi lain, ditemukan fenomena menarik terkait kelompok responden yang belum bekerja. Meskipun memiliki tingkat kompetensi dan IPK yang relatif setara dengan responden yang telah bekerja, mereka tidak memperoleh pekerjaan karena tidak melakukan pelamaran. Hal ini menunjukkan bahwa faktor proaktivitas memiliki pengaruh yang sangat signifikan dalam menentukan masa tunggu kerja.

Temuan ini memperlihatkan bahwa variabel perilaku (*behavioral features*) seperti inisiatif dan strategi pencarian kerja merupakan komponen penting yang sering kali tidak terakomodasi dalam model akademik konvensional. Dalam perspektif komputasional, hal ini menunjukkan perlunya memasukkan variabel non-akademik dalam model prediksi untuk meningkatkan akurasi.

Analisis lanjutan juga mengidentifikasi adanya pola non-linear antara IPK dan keberhasilan kerja. Responden dengan IPK tinggi tidak selalu memiliki tingkat keberhasilan yang lebih baik dibandingkan responden dengan IPK sedang. Hal ini mengindikasikan bahwa hubungan antara variabel akademik dan employability tidak bersifat linear, sehingga pendekatan model prediktif yang lebih kompleks seperti machine learning non-linear menjadi relevan untuk digunakan.

Selain itu, analisis keselarasan bidang kerja menunjukkan bahwa metode pembelajaran berbasis praktik memiliki pengaruh signifikan terhadap tingkat linieritas pekerjaan dengan bidang studi. Lulusan yang aktif dalam kegiatan magang, proyek, dan praktikum cenderung memperoleh pekerjaan yang lebih sesuai dengan kompetensi akademiknya.

Namun demikian, tidak ditemukan hubungan langsung antara tingkat keselarasan kerja dengan besaran pendapatan. Diskusi terhadap temuan ini menunjukkan bahwa pendapatan lebih dipengaruhi oleh faktor eksternal seperti sektor industri, lokasi kerja, dan skala perusahaan. Dengan demikian, keselarasan kerja lebih tepat digunakan sebagai indikator kualitas pendidikan, sedangkan pendapatan mencerminkan dinamika pasar tenaga kerja.

Secara keseluruhan, temuan dalam penelitian ini menunjukkan bahwa pendekatan berbasis data mampu mengungkap hubungan kompleks antar variabel dalam kesiapan kerja lulusan. Kesenjangan kompetensi, pengalaman praktis, dan proaktivitas terbukti menjadi faktor utama yang saling berinteraksi dalam menentukan keberhasilan lulusan di dunia kerja. Selain itu, hasil ini juga menegaskan pentingnya pendekatan berbasis *feature engineering* dalam mengoptimalkan analisis, terutama pada dataset dengan ukuran terbatas.

Deklarasi Kepentingan

Penulis menyatakan bahwa tidak terdapat konflik kepentingan dalam penelitian, penulisan, maupun publikasi artikel ini.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada rekan tim yang terlibat dalam penelitian ini. Penelitian ini tidak menerima pendanaan khusus dari lembaga pemerintah, komersial, maupun organisasi nirlaba..

Ketersediaan Data

Data tracer study yang digunakan dalam penelitian ini tersedia dari penulis korespondensi dan dapat diakses berdasarkan permintaan yang wajar dengan tetap memperhatikan aspek privasi responden.

Penggunaan Ai Dan Deklarasi Penggunaan Ai Generatif

Penulis menggunakan teknologi AI generatif untuk membantu penyusunan awal manuskrip, perbaikan tata bahasa, serta pengembangan struktur penulisan artikel. Seluruh hasil yang dihasilkan telah ditinjau, diverifikasi, dan menjadi tanggung jawab penuh penulis.

Kesimpulan

Penelitian ini menunjukkan bahwa pendekatan feature engineering dapat digunakan untuk menyusun indikator kuantitatif berbasis data tracer study dalam mengevaluasi kesiapan kerja lulusan. Hasil analisis mengidentifikasi adanya kesenjangan kompetensi pada aspek Teknologi Informasi dan Pengembangan Diri, serta menunjukkan bahwa pengalaman praktis memiliki keterkaitan deskriptif dengan status pekerjaan. Namun demikian, penelitian ini bersifat eksploratif dan tidak menggunakan metode pemodelan prediktif maupun uji statistik inferensial. Oleh karena itu, temuan yang diperoleh perlu ditindaklanjuti dengan penelitian lanjutan yang menggunakan pendekatan statistik atau machine learning untuk memperoleh validasi yang lebih kuat. Pendekatan ini diharapkan dapat menjadi dasar pengembangan sistem analitik tracer study di masa mendatang.

Daftar Pustaka

[1] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ. 2022 91*, vol. 9, no. 1, pp. 11-, Mar. 2022, doi: 10.1186/S40561-022-00192-Z.

[2] J. C. Agoylo, F. V Docena, M. Ann, C. Paulin, and K. N. Subang, "Career Outcomes And Employability Of It Graduates: A Tracer Study," *Wisdom J. Humanit. Soc. Sci.*, vol. 1, no. 5(November), pp. 40–48, Nov. 2024, Accessed: Apr. 02, 2026. [Online]. Available: https://so19.tci-thaijo.org/index.php/WJHS/article/view/1078.

[3] M. Andaya, J. R. S. Pedro, C. L. So, and B. Tanyag, "Employability of Computing Students in the Age of Disruption: A Graduate Tracer Study," *Eur. J. Educ. Pedagog.*, vol. 5, no. 1, pp. 82–87, Feb. 2024, doi: 10.24018/EJEDU.2024.5.1.793.

[4] G. Feng and M. Fan, "Research on learning behavior patterns from the perspective of educational data mining: Evaluation, prediction and visualization," *Expert Syst. Appl.*, vol. 237, p. 121555, Mar. 2024, doi: 10.1016/J.ESWA.2023.121555.

[5] M. U.-J. of P. S. Psychology and undefined 2022, "Graduates Employability Analysis using Classification Model: A Data Mining Approach.," *Res. UsitaJournal Posit. Sch. Psychol. 2022•researchgate.net*, Accessed: Apr. 02, 2026. [Online]. Available: https://www.researchgate.net/profile/Maricris-Usita/publication/393656953_Graduates_Employability_Analysis_using_Classification_Model_A_Data_Mining_Approach/links/687494343c12dc437a3e988b/Graduates-Employability-Analysis-using-Classification-Model-A-Data-Mining-Approach.pdf.

[6] M. A. Sapaat, A. Mustapha, J. Ahmad, K. Chamili, and R. Muhamad, "A Classification-Based Graduates Employability Model for Tracer Study by MOHE," *Commun. Comput. Inf. Sci.*, vol. 188 CCIS, no. PART 1, pp. 277–287, 2011, doi: 10.1007/978-3-642-22389-1_25.

[7] A. C. Albina and L. P. Sumagaysay, "Employability tracer study of Information Technology Education graduates from a state university in the Philippines," *Soc. Sci. Humanit. Open*, vol. 2, no. 1, p. 100055, Jan. 2020, doi: 10.1016/J.SSAHO.2020.100055.

[8] M. I. Ibrahim, "Career Pattern Analysis of SMKN 1 Stabat Graduates Using K-Means Clustering Algorithm on Tracer Study Dataset | Journal Of Data Science," 2025. https://ejournal.seaninstitute.or.id/index.php/visualization/article/view/6543 (accessed Apr. 02, 2026).

[9] R. Maqsood, P. Ceravolo, M. Ahmad, and M. S. Sarfraz, "Examining students' course trajectories using data mining and visualization approaches," *Int. J. Educ. Technol. High. Educ. 2023 201*, vol. 20, no. 1, pp. 55-, Oct. 2023, doi: 10.1186/S41239-023-00423-4.

[10] F. F. Abdulloh, M. Rahardi, A. Aminuddin, S. D. Anggita, and A. Y. A. Nugraha, "Observation of Imbalance Tracer Study Data for Graduates Employability Prediction in Indonesia," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 8, pp. 169–174, 2022, doi: 10.14569/IJACSA.2022.0130820.

[11] E. D. Lardizabal, "Tracer Study: Employability Criterion of the Teacher Education Program Graduates of PalSU-Quezon Campus," *Asian J. Educ. Soc. Stud.*, vol. 51, no. 5, pp. 583–598, May 2025, doi: 10.9734/AJESS/2025/V51I51942.