

Sistem Rekomendasi Pekerjaan dengan Graph Neural Network

Muhammad Syahrul Bachtiar, Wempy Aditya Wiryawan, Andika Nur Islamy

Jurusan Informatika, Fakultas Teknik/Universitas Muhammadiyah Malang rullbchtr2207, adityawempy, andikanurislamy {@webmail.umm.ac.id}

Abstrak

Penelitian ini mengembangkan Sistem Rekomendasi Pekerjaan berbasis Graph Neural Network (GNN) untuk mengatasi tantangan pencarian kerja dan rekrutmen di era digital. Dengan menggunakan data profil LinkedIn dan lowongan kerja, sistem merepresentasikan keterkaitan antara keterampilan, pengalaman, dan pekerjaan sebagai graph. Metode yang digunakan mencakup Graph Convolutional Network (GCN) untuk clustering pekerjaan dan GraphSAGE untuk prediksi hubungan antara profil dan lowongan. Hasil menunjukkan pengelompokan pekerjaan yang koheren serta performa rekomendasi moderat (AUC 0.6088, AP 0.5610)

Pendahuluan

Dalam era digital, pasar kerja semakin dinamis dengan persaingan yang ketat. Banyak pencari kerja mengalami kesulitan dalam menemukan pekerjaan yang sesuai dengan keterampilan dan pengalaman mereka, sementara perusahaan juga menghadapi tantangan dalam menemukan kandidat yang tepat.

1 Tujuan Penelitian

- Membantu pencari kerja menemukan pekerjaan
- Meningkatkan efisiensi proses perekrutan
- Memanfaatkan representasi berbasis graph untuk memahami dan memodelkan hubungan kompleks



Problem Statement

Bagaimana cara mengatasi kesulitan pencari kerja dan perusahaan dalam menemukan kecocokan antara keterampilan dan pengalaman pekerjaan?



menghasilkan rekomendasi pekerjaan yang lebih relevan dan akurat dibandingkan metode tradisional.

Dataset

Terdapat 2 dataset yang digunakan dalam penelitian ini:

- LinkedIn Professional Profiles Dataset: Merepresentasikan pencari kerja (node pengguna) beserta keterampilan dan pengalaman mereka.
- LinkedIn Job Postings Dataset : Merepresentasikan lowongan pekerjaan (node pekerjaan) beserta detailnya.

Total Records (DS 1) 1.000 rows

Total Records (DS 2) 5.000 rows

Features (DS 1) 26 columns

Features (DS 2) 31 columns

Data Challenges

Ekstraksi informasi keterampilan yang relevan secara konsisten dari beragam deskripsi pengguna dan pekerjaan menjadi tantangan utama dalam penyiapan dataset ini

Metodologi

Pendekatan metodologis yang digunakan dalam penelitian ini mengadopsi pipeline data science, meliputi pengumpulan data, preprocessing, feature ekstraksi, pemodelan, dan evaluasi.



Data Collection Pengumpulan data lowongan pekerjaan dan profile pekerja



Pembersihan data, ekstraksi dan profil pengguna



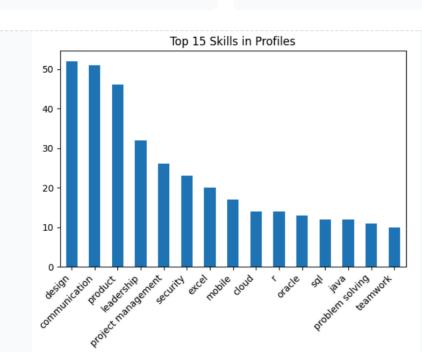
yang dihasilkan GCN untuk mengelompokkan pekerjaan



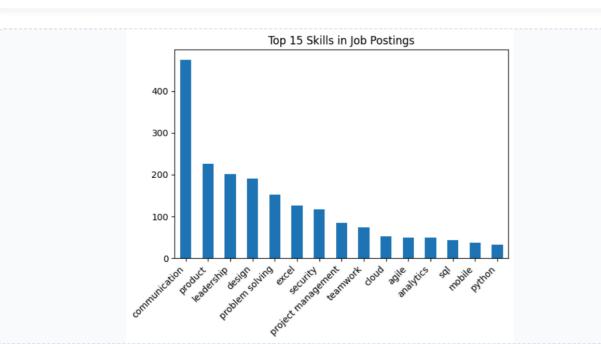
skill dari deskripsi pekerjaan



Link Prediction Membangun graph bipartite antara profil pencari kerja dan pekerjaan



Gambar 1: Distribusi 15 keterampilan teratas pengguna

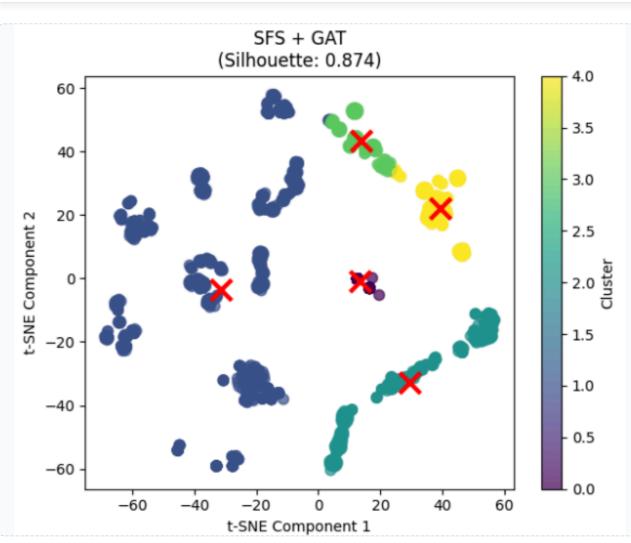


Gambar 2: Distribusi 15 keterampilan teratas pekerjaan

Hasil Eksperimen

Eksperimen menggunakan dua jenis model Graph Neural Network utama: Graph Convolutional Network (GCN) dan Graph Attention Network (GAT). Berikut hasil evaluasi dari berbagai skenario:

METODE	N_CLUSTERS	SILHOUETTE_SCORE	N_EDGES
Tanpa Feature Selection			
GNN	5	0.437164	1152352
GAT	5	0.484394	1152352
Chi-Square			
GNN	5	0.732855	1894124
GAT	5	0.757058	1894124
SFS			
GNN	5	0.801085	993108
GAT	5	0.873675	993108



Gambar 4: Visualisasi hasil dari clustering jenis lowongan pekerjaan

Analisis

Proses data mining yang telah dilakukan menggunakan GNN untuk clustering dan rekomendasi pekerjaan telah menghasilkan beberapa temuan penting. Interpretasi dari temuan ini dapat memberikan wawasan yang berharga.

Performa Model Terbaik

SFS + GAT menunjukkan performa terbaik dengan Silhouette 0.873675, mengindikasikan cluster yang padat & terpisah dengan baik.

Pengaruh Seleksi Fitur

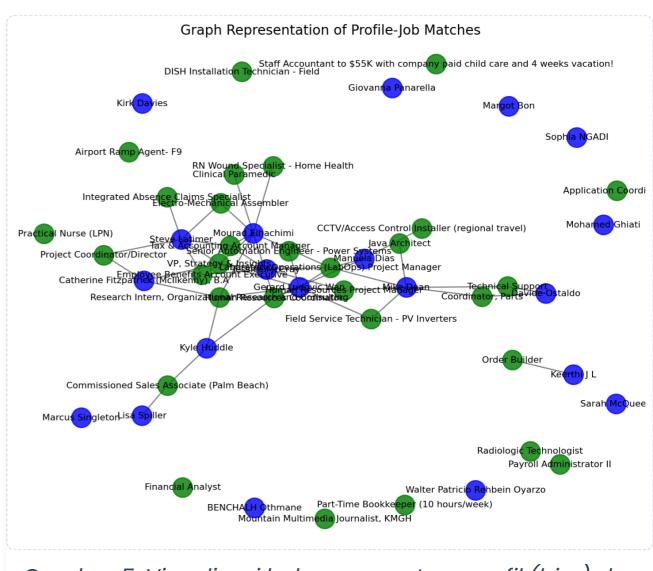
Seleksi fitur meningkatkan performa - SFS lebih unggul dari Chi-Square. Hal ini membantu model fokus di fitur penting & mengurangi noise.

Karakteristik Cluster

Analisis cluster menunjukkan pengelompokan alami: Cluster teknis, cluster soft skills, dan cluster penjualan/pemasaran.

Analisis Kesalahan

DBSCAN mengidentifikasi 4.11% noise points (66 dari 1605), terutama pekerjaan dengan kombinasi keterampilan unik / deskripsi ambigu.



Gambar 5: Visualisasi hubungan antara profil (biru) dan pekerjaan (hijau)

Kesimpulan

Penelitian ini berhasil mengimplementasikan GNN untuk sistem rekomendasi pekerjaan, dengan hasil pengelompokan pekerjaan koheren (Silhouette 0.874) dan model rekomendasi AUC 0.74040. Eksperimen menunjukkan bahwa kombinasi seleksi fitur (SFS) dan arsitektur GAT memberikan performa terbaik.

🛊 Kontribusi Utama

Penelitian ini menunjukkan efektivitas GNN, khususnya GAT, untuk analisis pasar kerja berbasis graf, dengan peningkatan kualitas clustering 100% melalui seleksi fitur SFS. Arsitektur hybrid menggabungkan GNN (clustering) dan GraphSAGE (rekomendasi), serta menganalisis metrik clustering dan rekomendasi secara komprehensif.

Future Work



https://doi.org/10.1109/bigdata.2017.8258088

User Feedback Feature Integrasi mekanisme

feedback pengguna untuk pembelajaran berkelanjutan



Mengeksplorasi arsitektur GNN yang lebih kompleks

Referensi

- A Comprehensive Survey of Artificial Intelligence Techniques for Talent Analytics. (2022). Arxiv.org. https://arxiv.org/html/2307.03195v2
- LightGNN: Simple Graph Neural Network for Recommendation. (2025). Arxiv.org. https://arxiv.org/html/2501.03228v1

• LinkSAGE: Optimizing Job Matching Using Graph Neural Networks. (2024). Arxiv.org. https://arxiv.org/html/2402.13430vl Shalaby, W., AlAila, B., Korayem, M., Pournajaf, L., AlJadda, K., Quinn, S., & Zadrozny, W. (2017). Help me find a job: A graph based approach for job recommendation at scale. 2017 IEEE International Conference on Big Data (Big Data).