

PROPOSAL PENELITIAN
SEMINAR BIDANG KAJIAN

Knowledge Graph pada Machine Learning untuk

Manajemen Talenta



Disusun Oleh :

Fikri Fadlillah

NIM : 99220703

PROGRAM DOKTOR TEKNOLOGI INFORMASI

UNIVERSITAS GUNADARMA

2021

Abstrak :

Fikri Fadlillah, S.T, MMSI

Knowledge Graph pada Machine Learning untuk Manajemen Talenta

Ditetapkannya Undang-Undang Nomor 5 Tahun 2014 tentang Aparatur Sipil Negara (ASN) menandai momentum peningkatan profesionalisme pada birokrasi pemerintahan di Indonesia. Salah satu aspek penting dalam pengelolaan ASN adalah menerapkan manajemen karir ASN. Kondisi saat ini menunjukkan lembaga pemerintahan belum memiliki prosedur sistematis yang mampu mencocokkan kompetensi dan kualifikasi suatu posisi dengan kompetensi dan kualifikasi pegawai. Berdasarkan hal di atas perlu dibangun suatu sistem yang lebih terencana dan sistematis untuk menerapkan manajemen ASN berbasis kompetensi mulai dari sistem rekrutmen, penempatan, pengembangan, dan promosi jabatan yang disebut sebagai manajemen talenta. Pada penelitian ini akan membahas mengenai pemetaan kompetensi dan kualifikasi jabatan dengan kompetensi dan kualifikasi yang dimiliki pegawai menggunakan *knowledge graph*. Metode Penelitian terdiri dari : (1) Akusisi data, (2) Pra-pemrosesan, (3) Pengenalan entitas bernama, (4) Menghubungkan entitas, dan (5) Ekstraksi relasi.

Daftar Pustaka (2007 - 2020)

Kata Kunci : *Machine Learning, Knowledge Graph, Manajemen Talenta*

BAB I

PENDAHULUAN

1.1. Latar Belakang

Ditetapkannya Undang-Undang Nomor 5 Tahun 2014 tentang Aparatur Sipil Negara (ASN) menandai momentum peningkatan profesionalisme pada birokrasi pemerintahan di Indonesia. Pemerintah Indonesia memulai langkah serius untuk mengimplementasikan sistem manajemen ASN untuk meningkatkan pegawai ASN guna memberikan pelayanan prima. UU ASN mendorong aspek kualifikasi, kompetensi dan kinerja menjadi isu utama sebagai fokus reformasi pelayanan publik di institusi pemerintah. Salah satu aspek penting dalam pengelolaan ASN adalah menerapkan manajemen karir ASN. Kondisi saat ini menunjukkan lembaga pemerintahan belum memiliki prosedur sistematis yang mampu mencocokkan kompetensi dan kualifikasi suatu posisi dengan kompetensi dan kualifikasi pegawai. Berdasarkan hal di atas perlu dibangun suatu sistem yang lebih terencana dan sistematis untuk menerapkan manajemen ASN berbasis kompetensi mulai dari sistem rekrutmen, penempatan, pengembangan, dan promosi jabatan yang disebut sebagai manajemen talenta. Manajemen talenta merupakan sistem untuk menarik, mengidentifikasi, mengembangkan, mempromosikan dan mempertahankan ASN yang memiliki potensi tinggi sebagai aset yang berharga bagi organisasi (Irfan, 2017). Berdasarkan uraian diatas, manajemen talenta membutuhkan algoritma yang digunakan untuk pemetaan kualifikasi dan kompetensi jabatan dengan kualifikasi dan kompetensi yang dimiliki pegawai, sehingga dilakukan penelitian mengenai “Knowledge Graph pada Machine Learning untuk Manajemen Talenta”.

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalah dalam penelitian ini terdiri dari :

- Bagaimana mengembangkan metode pemetaan kualifikasi dan kompetensi jabatan dengan kualifikasi dan kompetensi pegawai dengan knowledge graph ?
- Bagaimana mengoptimalkan operasional manajemen talenta pada instansi ?

1.3. Batasan Masalah

Berdasarkan rumusan masalah, batasan masalah pada penelitian ini terdiri dari :

- Pemetaan kompetensi dan kualifikasi pegawai ke dalam knowledge graph

- Konstruksi knowledge graph melalui tahapan akusisi data, identifikasi entitas, menghubungkan entitas, dan ekstraksi relasi

1.4. Tujuan Penelitian

Tujuan khusus dari penelitian ini adalah menghasilkan metode pemetaan kualifikasi dan kompetensi jabatan dengan kualifikasi dan kompetensi pegawai. Sedangkan tujuan umum dari penelitian ini adalah mengoptimalkan operasional manajemen talenta pada instansi.

1.5 Kontribusi Penelitian

Kontribusi penelitian secara keilmuan adalah menemukan cara baru/penambahan/modifikasi dari metode yang ada untuk pemetaan kualifikasi dan kompetensi jabatan dengan kualifikasi dan kompetensi pegawai dengan *knowledge graph*. Sedangkan kontribusi penelitian secara teknologi adalah menghasilkan perangkat lunak yang dapat membantu dan memudahkan pengguna dalam melakukan pemetaan kualifikasi dan kompetensi jabatan dengan kualifikasi kompetensi pegawai.

BAB II

TINJAUAN PUSTAKA

Bab ini menguraikan tentang studi literatur terkait dengan *knowledge graph* pada *machine learning* pada manajemen talenta.

2.1. Pengenalan Manajemen Talenta

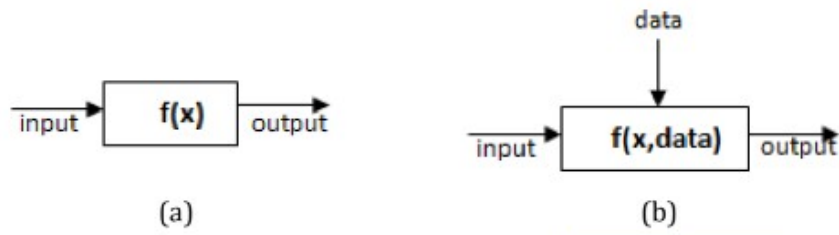
Menurut Amstrong (dalam Erni T.S, 2016), manajemen talenta adalah seperangkat aktivitas yang terintegrasi dan komprehensif yang bertujuan mengamankan aliran talenta dalam organisasi dan membentuk pemikiran bahwa talenta adalah sumber daya utama dalam organisasi.

Kemudian Gaspers (dalam Erni T.S, 2016) mendefinisikan bahwa manajemen talenta adalah proses menahan sumber daya manusia yang terkait dengan proses mengembangkan dan memperkuat karyawan baru, memelihara dan mengembangkan karyawan yang sudah ada, serta menarik sebanyak mungkin karyawan yang memiliki kompetensi, komitmen dan karakter agar mau bekerja dalam perusahaan.

Sedangkan Baum (dalam Erni T.S, 2016) menyatakan manajemen talenta sebagai mindset organisasi dalam menjamin ketersediaan talenta yang merupakan orang yang tepat dengan pekerjaan yang tepat. Sehingga manajemen talenta adalah serangkaian aktivitas terintegrasi dalam mengelola karyawan berkinerja tinggi di semua level organisasi dengan komponen utama adalah menarik talenta melalui pengadaan dan orientasi, mengembangkan talenta melalui manajemen kinerja, review talenta, serta mempertahankan talenta melalui perencanaan karier, suksesi dan mengikat talenta menurut Wahyuningtyas (dalam Erni T.S, 2016)

2.2. Machine Learning

Machine learning adalah aplikasi atau bagian dari kecerdasan buatan yang membuat sistem memiliki kemampuan belajar secara otomatis dan meningkatkan kemampuannya berdasarkan pengalaman tanpa diprogram secara eksplisit yang dimana program komputer tidak ditulis secara statis. Fokus *machine learning* terdapat pada pengembangan program komputer yang dapat mengakses data dan belajar dari data tersebut (Purba, 2017) Ilustrasi machine learning terdapat pada gambar 2.1.



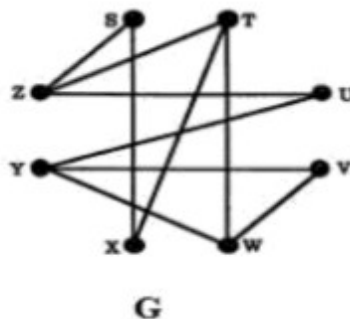
Gambar 2.1 Diagram Algoritma Umum dan *Machine Learning* (Purba, 2017)

Berdasarkan gambar 2.1 (a), terdapat suatu proses dimana output merupakan keluaran fungsi $f(x)$ yang dimana x adalah masukan fungsi. Sedangkan pada gambar 2.1(b), output merupakan keluaran fungsi f dimana yang menjadi masukan adalah input dan data.

2.3. Definisi Graf

Menurut Farida (2019), graf adalah diagram untuk merepresentasikan objek diskrit dan hubungan antara objek-objek tersebut. Suatu Graf terdiri dari 2 himpunan berhingga, yaitu:

- $V(G)$, adalah himpunan verteks atau simpul, yang dinotasikan dengan huruf V untuk himpunan tak kosong dari simpul- simpul
- $E(G)$, adalah himpunan sisi (*edge*), dinotasikan dengan huruf E , untuk menyatakan himpunan dari sisi yang menghubungkan sepasang simpul. Setiap elemen e dari $E(G)$ merupakan sebuah pasangan tak berurutan dari simpul-simpul di $V(G)$. Contoh Graf terdapat pada gambar 2..



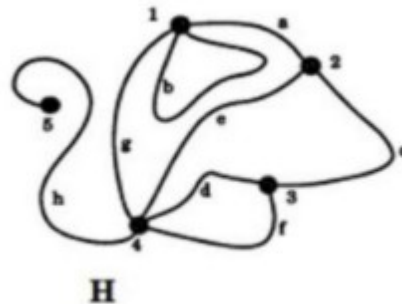
Gambar 2 Graf G dengan 8 simpul dan 10 sisi (Farida, 2019)

Berdasarkan gambar diatas, himpunan verteks dan himpunan sisi graf G adalah :

$$V = \{S, T, U, V, W, X, Y, Z\}$$

$$E = \{(S,X),(S,Z) ,(T,W) ,(T,X) ,(T,Z) ,(U,Y) ,(V,W) ,(V,Y) ,(W,Y)\}$$

Verteks atau simpul saring disebut titik / noktah. Jika e adlah sisi dengan ujung u dan v maka e dikatakan sisi yang menghubungkan u dan v . Sehingga terdapat kemungkinan bahwa sebuah sisi memiliki simpul-simpul ujung yang sama atau sebuah simpul terhubung dengan dirinya sendiri. Sisi yang demikian disebut *loop*. Contoh graf yang terdapat loop terdapat pada gambar 3.



Gambar 3 Graf H dengan 6 simpul dan 8 sisi (Farida, 2019)

Berdasarkan gambar diatas, terdapat *loop* pada sisi b pada graf H yang menghubungkan simpul 1 dengan dirinya sendiri. Kemudian sisi tidak selalu digambarkan sebagai sebuah garis lurus, tetapi dapat digambarkan berupa garis melengkung.

2.4. Knowledge Graph (KG)

Knowledge Graph adalah metode yang digunakan untuk menganalisis teks dan mempresentasikannya ke dalam bentuk graf (Zhang dan Hoede, 2000). Menurut Zhang (2002), *knowledge graph* adalah pendekatan baru untuk menyatakan manusia yang lebih

memfokuskan pada aspek semantik. *Knowledge Graph* terdiri atas *concept* (*token*, *type*, dan *name*) dan *relations*.

a) Concept

Concept merupakan komponen terpendng dalam pemikiran manusia dan mampu menjadi prosedur dalam membentuk suatu pengertian dari khusus ke umum atau sebaliknya (Zhang, dalam Mahmuda:2010) *Concept* dalam KG dapat dinyatakan sebagai *token*, *name*, dan *type*.

- *Token* merupakan *concept* yang dipahami oleh seseorang menurut cara pandangnya.
- *Type* adalah *concept* yang berupa informasi umum,
- *Name* adalah sesuatu yang bersifat individual.

b) Relations

Relations pada *knowledge graph* terdiri dari 9 *binary relations* yang terdiri dari :

- Relasi *Alikeness (ALI)*
Relasi ALI digunakan untuk menghubungkan sebuah *type* dengan *token* (Zhang, dalam Mahmuda:2010)
- Relasi *Causality (CAU)*
Relasi CAU digunakan untuk menghubungkan dua *token* yang memiliki sebab akibat (Zhang, dalam Mahmuda:2010)
- Relasi *Equality (EQU)*
Relasi EQU digunakan untuk menghubungkan sebuah *name* dengan *token* (Berg, 1993).
- Relasi *Subset (SUB)*
Relasi SUB digunakan untuk menghubungkan kedua *token* yang menyatakan dua rangkaian yang saling bertautan (Zhang, dalam Mahmuda:2010)
- Relasi *Disparatness (DIS)*
Relasi DIS digunakan untuk menghubungkan dua *token* yang tidak mempunyai satu elemen yang sama dengan yang lainnya (Zhang, dalam Mahmuda:2010)
- Relasi *Attribution (PAR)*

Relasi PAR digunakan untuk menjelaskan bahwa satu elemen berkaitan dan memiliki sifat elemen lainnya (Zhang, dalam Mahmuda:2010)

- Relasi *Ordering (ORD)*

Relasi ORD menyatakan bahwa dua benda memiliki urutan tertentu satu sama lain, baik urutan waktu atau tempat. (Zhang, dalam Mahmuda:2010)

- Relasi *Skolem (SKO)*

Menurut Berga (1993), relasi SKO menyatakan informasi bergantung dan mampu menggambarkan kuantifikasi. (Zhang, dalam Mahmuda:2010)

- *Fokus (F)*

F digunakan untuk menunjukkan fokus dari suatu graph (Nurdiati dan Hoede, 2009)

2.5. Penelitian Terdahulu

Sebagai bahan perbandingan dan kajian, maka penelitian yang akan dilakukan tidak terlepas dari hasil penelitian-penelitian terdahulu. Hasil penelitian sebelumnya yang menjadi perbandingan berhubungan dengan topik penelitian, yaitu mengenai pemodelan pengetahuan menggunakan *Knowledge Graph*

Penelitian Mahmuda (2007), yang melakukan penelitian *knowledge graph* untuk konstruksi pola *Word Graph* Frasa Kata dengan alur penelitian : (1) identifikasi frasa kata Bahasa Indonesia, (2) implementasi relasi dalam bentuk *word graph* frasa kata, (3) pengelompokkan frasa kata berdasarkan bentuk *word graph* , (4) membuat aturan pembentukan *word graph* frasa kata, (5) menguji aturan *word graph* frasa kata. Dari penelitian tersebut diperoleh aturan pembentukan *word graph* frasa kata pada Bahasa Indonesia, dan dapat diterapkan untuk semua frasa kata yang sama strukturnya

Penelitian Wulan (2010), yang melakukan penelitian pembentukan *word graph* preposisi Bahasa Indonesia dengan *Knowledge Graph*, dengan metode mengekspresikan arti dari sebuah kata dengan *word graph*. Berdasarkan penelitian tersebut, diperoleh hasil bahwa preposisi punya makna majemuk, sehingga bentuk *word graph* untuk satu preposisi data diimplementasikan ke dalam bentuk *word graph* yang berbeda, tetapi terdapat juga preposisi bermakna tunggal

Penelitian dari Lukman (2020), tentang penggunaan *Knowledge Graph* untuk visualisasi tematik Al-Qur'an menggunakan basis data grafik Neo4j, dengan data tematik Al-Quran diambil dari Al-Qur'an Amazing dan data Al-Qur'an diambil dari

qurandatabase.org, diperoleh hasil visualisasi hubungan ayat-ayat Al-Quran dengan tematiknya dengan nilai rata-rata precision 1 dan F-score 0.56 untuk pencarian berdasarkan surat dan ayat dan nilai rata-rata precision 1 dan F-score 1 untuk pencarian berdasarkan tema.

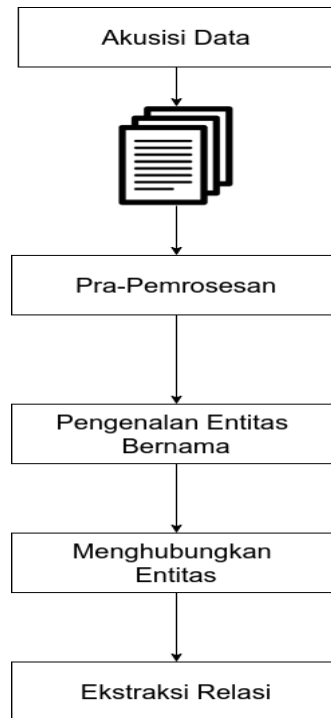
Penelitian yang dilakukan oleh Ghanibi, M. et. al. (2020), menciptakan *novel software tool* bernama FoodKG, yang memperkaya *knowledge graphs* mengenai makanan, energi, dan air menggunakan teknik *machine learning* lanjutan dengan tujuan untuk meningkatkan pengambilan keputusan dan penjelajahan pengetahuan dalam bidang makanan, energi, dan air. Metodologi penelitian yang dilakukan diantaranya menggunakan AGOROVEC untuk memproduksi *domain-specific embedding model*, ekstraksi entitas dengan teknik *Natural Language Processing* (NLP), klasifikasi teks, mengukur similaritas semantik diantara dua *term / concept*, menentukan *scientific terms*, dan prediksi relasi diantara *concept* yang berbeda. Hasil penelitian menunjukkan bahwa AGROVEC yang menggunakan algoritma *Spearman Correlation Coefficient* dapat mengungguli model terlatih lainnya dalam set data *graph* yang sama.

Penelitian Liu L. et. al. (2020), mengusulkan metode *task of job posting generation* untuk menciptakan kebutuhan kualifikasi dari lamaran pekerjaan berdasarkan kondisi yang diberikan dari deskripsi pekerjaan dan informasi dasar perusahaan. Metode penelitian terdiri dari akuisisi data dari situs rekrutmen yang terkenal di china sebanyak 107,616 posting lamaran pekerjaan dari 6 domain industri yang berbeda, melakukan *encoding* menggunakan *bi-directional Long Short Term Memory* untuk memperoleh representasi vektor mengenai ketergantungan *long-term*, memprediksi *skill* dengan *Multi Label Classification* (MLC) *task*, menyempurnakan *skill* dari kebutuhan pekerjaan yang akan diciptakan, menciptakan kebutuhan pekerjaan, dan *Training* serta pengambilan kesimpulan.

BAB III

METODOLOGI PENELITIAN

Secara garis besar, metodologi penelitian diilustrasikan pada gambar 3.1.



Gambar 3.1. Metodologi Penelitian

3.1. Akusisi Data

Akuisisi data adalah tahap pengumpulan data yang dibutuhkan dalam penelitian. Dalam penelitian ini, data dikumpulkan dari database yang digunakan untuk menyimpan data kualifikasi dan kompetensi pegawai. Data di ambil dari database melalui *Application Programming Interface* yang telah di sediakan, kemudian data di konversi ke dalam file berekstensi *.csv

3.2. Pra-pemrosesan

Pada tahap pra-pemrosesan, dilakukan penerapan teknik pada teks input dari tahapan akusisi data , untuk menguraikan teks menjadi istilah dan simbol atomik. Untuk kebutuhan pra-pemrosesan, menggunakan Natural Language Processing

3.3. Pengenalan Entitas Bersama

Pada tahap ini, dilakukan identifikasi penyebutan entitas bernama dalam teks, untuk penyebutan orang, organisasi, lokasi, dan kemungkinan jenis lainnya. Teknik ini menggunakan pendekatan moderen dengan memanfaatkan fitur leksikal.

3.4. Menghubungkan Entitas

Pada tahap ini, suatu entitas di dalam teks akan diidentifikasi berdasarkan *knowledge graph* dari sumber lain yang telah ada.

3.5. Ekstraksi Relasi

Pada tahap ini, dilakukan ekstraksi relasi diantara entitas yang terdapat dalam teks, sehingga terbentuk *knowledge graph*.

DAFTAR PUSTAKA

- Cappelli, Peter. *Talent on Demand*. Metode Baru Mendapatkan SDM Bertalenta Tepat Jumlah, Tepat Kualifikasi, dan Tepat Waktu. Diterjemahkan oleh Wendra Triana Rafelina. Hal 10-12. 2009
- Daru K., Purba. *Machine Learning* : Teori, Program Dan Studi kasus, hal. 20–21, Sleman , Deepublish, 2020.
- Mahmuda. Konstruksi Pola *Word Graph* Frasa Kata Menggunakan Metode *Knowledge Graph*. Thesis Institut Pertanian Bogor. 2010.
- Erni T.S. dan Ratri W. Manajemen Talenta Terintegrasi , hal. 3 – 5, Yogyakarta, Andi Offset, 2016.
- Anggraeni, Wulan. Pembentukan *Word Graph* Preposisi Bahasa Indonesia Menggunakan Metode Knowledge Graph. Jurnal Ilmiah Faktor Exacta 2010, Vol. 3 No.2.
- Huliyah, Khodijah. Rekayasa Memahami Teks Menggunakan Metode *Knowledge Graph*. Thesis Institut Pertanian Bogor. 2007.
- Nul Hakim, Lukman. Visualisasi Tematik Al-Quran Berbasis *Knowledge Graph*. Jurnal Lingustik Komputasional. Vol. 3 No. 1. 2020.
- Gharibi M., Zachariah A., dan Rao P. A Tool to Enrich Knowledge Graphs Using Machine Learning Techniques. *Frontiers In Big Data*. Vol. 3 No. 12. 2020
- Liu, Liting., et. al. Hiring Now: A Skill-Aware Multi-Attention Model for Job Posting Generation. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Vol. 3 No. 12. hal. 3096 – 3104. 2020