

***COSINE SIMILARITY WITH WORD EMBEDDING
FOR LEARNING DATA ANALYTICS DASHBOARD
USING GLOVE***

LAPORAN INTERNSHIP I

Laporan ini dibuat untuk memenuhi persyaratan kelulusan
mata kuliah Program Internship I



Dibuat Oleh,

1.17.4021 Muhammad Fahmi

PROGRAM DIPLOMA IV TEKNIK INFORMATIKA

POLITEKNIK POS INDONESIA

BANDUNG

2020

***COSINE SIMILARITY WITH WORD EMBEDDING
FOR LEARNING DATA ANALYTICS DASHBOARD
USING GLOVE***

INTERNSHIP PROGRAM REPORT I

This report is prepared to fulfill graduation requirements
Internship program courses



Created by,

1.17.4021 Muhammad Fahmi

**INFORMATICS ENGINEERING DIPLOMA IV PROGRAM
POLITEKNIK POS INDONESIA
BANDUNG
2020**

ABSTRAK

Data merupakan kumpulan informasi yang diperoleh dari beberapa sumber untuk disajikan dalam bisnis dunia. Ilmu data adalah teknologi untuk memecahkan masalah analitik yang kompleks. Dalam ilmu data, ada juga analisis prediksi data yang akan disaring dan menghasilkan data yang akurat sesuai data aktual. *Learning Data Analytics* akan membuat daftar rekomendasi pelatihan yang akan dilaksanakan di Telkom Corporate University Bandung sesuai dengan judulnya *Cosine Similarity With Word Embedding For Learning Data Analytics Dashboard Using GloVe*. Seluruh alur kerja akan menggunakan bahasa pemrograman python. Proses rekomendasi dimulai dengan mengumpulkan sumber data untuk kursus atau pelatihan online yang diadakan secara online; menggunakan *scraping* atau *crawling*. Penulis akan mengambil topik, judul, deskripsi, dan lain-lain dalam kursus online. Kemudian dilakukan penyematan kata menggunakan *GloVe* dan *cosine similarity*, dengan fokus pada data pembelajaran dan tujuan yang sudah ada di Telkom Corporate University Bandung. Di sini penulis akan menghasilkan rekomendasi pelatihan yang cocok diadakan di Telkom Corporate University sesuai data *subskill* dan *objective* yang sudah ada. Tidak lupa juga merekomendasikan saran pengembangan topik di tahun 2020 untuk dibandingkan dengan data tersebut untuk mendapatkan hasil rekomendasi pelatihan yang lebih akurat.

Kata Kunci : *GloVe, Learning Data Analytics, Word Embedding, Cosine Similarity*

ABSTRACT

Data is a collection of information obtained from several sources to be presented in the business world. Data science is a technology for solving complex analytical problems. There is also a predictive analysis of data in data science that will be filtered and the result of the analysis is accurate data according to the actual data. Learning Data Analytics will list training recommendations to be carried out at Telkom Corporate University Bandung according to the title Cosine Similarity With Word Embedding For Learning Data Analytics Dashboard Using GloVe. The entire workflow will use the python programming language. The recommendation process begins by collecting data sources for online courses or training held online, using scraping or crawling. The author will take the topic, title, description, etc., in the online course. Then do the embedding of words using GloVe and cosine similarity, focusing on learning data and objectives that already exist at Telkom Corporate University Bandung. Here, the writer will produce training recommendations that are suitable to be held at Telkom Corporate University according to existing subskill and objective data. Do not forget to recommend topic development suggestions in 2020 to be compared with this data to get more accurate training recommendation results.

Keyword : *GloVe, Learning Data Analytics, Word Embedding, Cosine Similarity*

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Puji syukur penulis panjatkan ke hadirat Allah SWT karena hanya dengan rahmat dan hidayahnya, laporan internship I ini dapat terselesaikan tanpa halangan berarti. Keberhasilan dalam menyusun laporan internship I ini tidak lepas dari bantuan berbagai pihak yang mana dengan tulus dan ikhlas memberikan masukan guna sempurnanya laporan internship I ini. Oleh karena itu, dalam kesempatan ini, dengan kerendahan hati penulis mengucapkan terima kasih kepada:

1. M. Yusril Helmi Setyawan, S.Kom., M.Kom. Selaku Ketua Program Studi DIV Teknik Informatika Politeknik Pos Indonesia,
2. Cahyo Prianto, S.Pd., M.T. Selaku Koordinator internship I tahun 2020/2021,
3. Orang Tua, yang selalu memberi dukungan dan doa,
4. Teman - teman yang juga memberi dukungan dan doa.

Penulis menyadari bahwa penyusunan laporan internship I ini jauh dari sempurna. Kritik dan saran dapat ditujukan kontak penulis. Dengan terselesaikannya laporan internship I, penulis berharap semoga laporan ini dapat menimbulkan manfaat bagi para pembaca.

Wassalamu'alaikum Wr. Wb.

Asahan, 05 Januari 2021

Penulis

DAFTAR ISI

ABSTRAK.....	i
<i>ABSTRACT</i>	ii
KATA PENGANTAR.....	iii
DAFTAR ISI.....	iv
DAFTAR GAMBAR.....	vi
DAFTAR TABEL.....	vii
DAFTAR SIMBOL.....	viii
DAFTAR SINGKATAN.....	ix
DAFTAR LAMPIRAN.....	x
BAB I PENDAHULUAN	I-1
1.1 Latar Belakang.....	I-1
1.2 Identifikasi Masalah.....	I-3
1.3 Tujuan dan Manfaat.....	I-3
1.4 Ruang Lingkup.....	I-3
1.5 Penelitian Sebelumnya.....	I-4
1.6 Sistematika Penulisan.....	I-5
BAB II LANDASAN TEORI	II-1
2.1 <i>Data Science</i>	II-1
2.2 <i>Python</i>	II-1
2.3 <i>Glove</i>	II-1
2.4 <i>Scraping</i>	II-2
2.5 <i>Recommendation System</i>	II-2
2.6 <i>Pandas</i>	II-2
2.7 Rekomendasi Pelatihan	II-2
2.8 <i>Word Embedding</i>	II-3
2.9 <i>Text Matching</i>	II-3
2.10 <i>Cosine Similarity</i>	II-3
2.11 <i>Natural Language Processing</i>	II-3

BAB III ANALISIS ORGANISASI PERUSAHAAN.....	III-1
3.1 Sejarah Perusahaan.....	III-1
3.2 Visi dan Misi Perusahaan.....	III-2
3.3 Strategi Perusahaan.....	III-3
3.4 Struktur Organisasi.....	III-3
3.4.1 <i>Job Description</i> Peserta Internship.....	III-3
3.5 Ruang Lingkup Internship.....	III-4
BAB IV METODOLOGI PENELITIAN.....	IV-1
4.1 Diagram Alur Metodologi Penelitian.....	IV-1
4.2 Tahapan – Tahapan Diagram Alur Metodologi Penelitian.....	IV-2
4.2.1 <i>Business Understanding</i>	IV-2
4.2.2 <i>Data Understanding</i>	IV-2
4.2.3 <i>Data Preparation</i>	IV-2
4.2.4 <i>Modelling</i>	IV-2
4.2.5 <i>Evaluation</i>	IV-3
BAB V ANALISIS DAN PERANCANGAN SISTEM.....	V-1
5.1 Analisis Sistem yang akan dibangun	V-1
5.1.1 Analisis Prosedur yang akan dibangun	V-1
5.1.1.1 <i>Flowchart</i> pada <i>Training Recommendation</i>	V-1
5.1.1.2 <i>Flowchart</i> pada <i>Expert Recommendation</i>	V-2
5.2 Perancangan Basis Data/Database Sistem.....	V-3
5.3 Perancangan User Interface Sistem.....	V-4
5.3.1 Perancangan Tampilan Home.....	V-4
5.3.2 Tampilan Menu <i>Training Recommendation</i>	V-4
5.3.3 Tampilan Menu <i>Emerging Topic</i>	V-5
5.4 Perancangan Arsitektur Perangkat Lunak dan Perangkat Keras..	V-5
5.4.1 Kebutuhan Perangkat Lunak.....	V-5
5.4.2 Kebutuhan Perangkat Keras.....	V-6

BAB VI PENGKAJIAN DAN EVALUASI.....	VI-1
6.1 Hasil Secara Keseluruhan	VI-1
6.2 Kemanfaatan Fungsi Sistem Terhadap Permasalahan	VI-2
BAB VII PENUTUP.....	VII-1
7.1 Kesimpulan dan Saran.....	VII-1
7.2.2 Kesimpulan.....	VII-1
7.2.3 Saran.....	VII-1
7.2 Lampiran-lampiran.....	VII-1
7.2.1 Laporan Kemajuan Bimbingan (Kambing).....	VII-2
7.2.2 Penilaian Dari Pembimbing Eksternal.....	VII-3
7.2.3 Curriculum Vitae dalam bahasa Inggris dan Indonesia.....	VII-4
DAFTAR PUSTAKA.....	xi

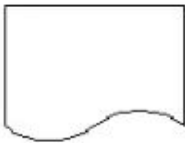
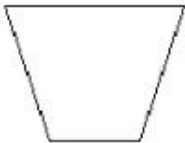

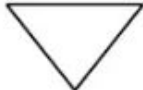

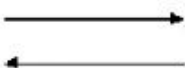

DAFTAR GAMBAR

Gambar 3.1 Logo Telkom.....	III-1
Gambar 3.2 Struktur Organisasi DDB Telkom.....	III-3
Gambar 4.1 Diagram Alur Penelitian	IV-1
Gambar 5.1 <i>Flowchart Training Recommendation</i>	V-1
Gambar 5.2 <i>Flowchart Expert Recommendation</i>	V-2
Gambar 5.3 Perancangan <i>Database</i>	V-3
Gambar 5.4 Perancangan Tampilan <i>Home</i>	V-4
Gambar 5.5 Perancangan Tampilan <i>Menu Training Recommendation</i>	V-4
Gambar 5.6 Perancangan Tampilan <i>Menu Emerging Topic</i>	V-5
Gambar 6.1 <i>Filter</i> untuk <i>Recommendation System</i>	VI-1
Gambar 6.2 Hasil <i>Expert Recommendation</i>	VI-1
Gambar 6.3 Hasil <i>Training Recommendation</i>	VI-2

DAFTAR TABEL

Tabel 1.1 Penelitian Sebelumnya.....	I-4
Tabel 5.1 Kebutuhan Perangkat Lunak.....	V-5
Tabel 5.1 Kebutuhan Perangkat Keras.....	V-6

DAFTAR SIMBOL

NO	SIMBOL	NAMA	KETERANGAN
1		Dokumen	I/O dalam format yang dicetak
2		Manual Operation	Proses yang terjadi di dalam flowmap
3		Proses komputerisasi	Merepresentasikan Input data atau Output data yang diproses atau Informasi
4		Arsip manual	Penyimpanan yang dapat diakses oleh komputer secara langsung
5		Decision	Menunjukkan pilihan keputusan
6		Aliran	Menunjukkan data-data yang mengalir pada sistem
7		Direct Data	Merupakan penyimpanan database

DAFTAR SINGKATAN

<i>NLP</i>	: <i>Natural Language Processing</i>
<i>Glove</i>	: <i>Global Vectors for Word Representation</i>
<i>CRISP-DM</i>	: <i>Cross-Industry Standard Process for Data Mining</i>

DAFTAR LAMPIRAN

Lampiran 1. Laporan Kemajuan Bimbingan (Kambing).....	VII-2
Lampiran 2. Penilaian Dari Pembimbing Eksternal.....	VII-3
Lampiran 3. Curriculum Vitae dalam bahasa Inggris dan Indonesia.....	VII-4

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data adalah kumpulan sebuah informasi atau deskripsi tentang sesuatu yang diperoleh dengan mengamati atau mencari sumber tertentu. Data yang diperoleh dapat berupa asumsi atau fakta karena belum diolah lebih lanjut. Setelah diolah dengan penelitian atau eksperimen, data dapat berbentuk lebih kompleks seperti *database*, informasi atau bahkan solusi untuk menyelesaikan masalah tertentu. Data juga diperlukan untuk diolah dan disajikan dalam dunia bisnis [1].

Dalam era sekarang ini muncul sebuah bidang profesi baru yaitu *data science*. *Data science* merupakan gabungan dari inferensi data, pengembangan algoritma dan juga teknologi untuk memecahkan masalah analitik yang kompleks. Dalam *data science* juga terdapat analisis prediktif suatu data untuk difilter dan ditemukan data yang benar agar menghasilkan suatu data yang akurat sesuai dengan data yang sebenarnya. *Data science* atau ilmu data adalah seperangkat prinsip fundamental yang mendukung dan memandu ekstraksi berprinsip informasi dan pengetahuan dari data. Ada ratusan jenis metode atau algoritma yang ada pada *data science* salah satunya yang ingin penulis jelaskan di proposal ini adalah *Recommendation system* [2].

Recommendation system atau Sistem rekomendasi biasanya melibatkan pemanfaatan hubungan antara fitur dan konten yang diketahui yang mendeskripsikan item (*content-based filtering*) atau pengguna serupa yang saling berinteraksi atau memberi peringkat item target (*collaborative filtering*) [3].

Recommendation system atau sistem rekomendasi merupakan bagian dari filtering pada data yang melakukan prediksi untuk suatu tujuan masalah. *Recommendation system* akan menghasilkan suatu daftar rekomendasi yang dapat memecahkan suatu masalah, dalam hal ini sesuai dengan judul yaitu Penerapan Metode *Recommendation System Based On Natural Language Processing* Dalam Rekomendasi Pelatihan (*Learning Data Analytics Telkom Corporate University*).

Pada judul tersebut penulis akan memakai bahasa pemrograman *python* dengan melibatkan beberapa *library* NLP seperti *pandas*, *csv*, *word2vec*, *glove* dan yang lainnya yang akan penulis jelaskan pada bab selanjutnya. Adapun output dari proposal yang akan dikerjakan oleh penulis adalah sebuah daftar rekomendasi pelatihan apa yang cocok untuk diadakan oleh *Telkom Corporate University*. Pada sistem rekomendasi penulis menggunakan *content-based filtering* atau *content based recommendation*. *Content Based Recommendation* memanfaatkan informasi beberapa item/ data untuk digunakan sebagai referensi yang terkait dengan informasi yang digunakan sebelumnya. Tujuan dari *content based recommendation* agar dapat memprediksi persamaan dari sejumlah informasi yang didapat dari data sebelumnya [4].

Pada kasus ini, sebuah informasi yang telah didapatkan dari *Telkom Corporate University* akan dimanfaatkan semaksimal mungkin untuk digunakan pada pengolahan dan analisis data. Data yang ada juga berbasis text yaitu NLP atau *Natural Language Processing* yang artinya bagaimana suatu data text yang berbahasa indonesia atau inggris dapat dimengerti oleh sistem komputer. Dalam pembuatannya, *content based filtering* menggunakan konsep perhitungan vektor, *TF-IDF*, dan *Cosine Similarity* yang intinya dikonversikan dari data / text menjadi berbentuk vektor.

Proses rekomendasi dimulai dari mengumpulkan sumber-sumber data *course* atau pelatihan yang ada diadakan secara online, dengan *scraping* atau *crawling* penulis akan mengambil topik, judul, deskripsi dan semua hal yang ada pada *course* online tersebut. Kemudian, akan dilakukan *word embedding* dan juga *matching* atau *cosine similarity* dengan data *learning focus* dan *objective* yang sudah ada pada *Telkom Corporate University*, disini akan menghasilkan suatu *text matching* dengan ini kita bisa melihat rekomendasi pelatihan apa yang cocok. Tidak lupa juga penulis membuat rekomendasi topik dan *emergic topic* yang trending di 2020 agar bisa disamakan dengan yang ada pada data agar mendapatkan hasil yang lebih akurat untuk rekomendasi pelatihan.

1.2 Identifikasi Masalah

Berdasarkan latar belakang di atas ada beberapa rumusan masalah, di antara lain ialah :

1. Bagaimana cara menerapkan metode *recommendation system* menggunakan *GloVe* pada proses rekomendasi pelatihan ?
2. Bagaimana cara mengolah dan menganalisis data untuk keperluan *recommendation system* ?
3. Bagaimana proses *Word Embedding* dapat menjadi suatu cara untuk memperoleh daftar rekomendasi pelatihan yang akurat ?

1.3 Tujuan dan Manfaat

Adapun tujuan dan manfaat pada proposal internship I ini adalah :

1. Menerapkan model *Word Embedding GloVe*.
2. Melakukan *scraping* data *course* dan topik tentang pelatihan yang trending serta mengumpulkan data-data yang sebelumnya.
3. Menerapkan *word embedding* atau text menjadi sebuah vektor dan juga menerapkan *model cosine similarity*.

1.4 Ruang Lingkup

Berdasarkan identifikasi masalah serta tujuan dan manfaat diatas, maka ruang lingkup atau batas dalam proses pengerjaan internship I ini adalah :

1. Metode *recommendation system* yang digunakan hanya *Word Embedding Cosine Similarity*.
2. Proses mengubah text menjadi vektor hanya memakai model *GloVe*.
3. Sistem rekomendasi yang dihasilkan hanya disajikan dan digunakan oleh *Telkom Corporate University*.

1.5 Penelitian Sebelumnya

Pada proposal ini, penulis mereferensikan beberapa penelitian sebelumnya yang serupa dengan judul internship I di proposal ini, di antara lain adalah :

Tabel 1.1: Penelitian sebelumnya

No	Judul	Kutip	Kesimpulan
1	<i>Content-based Recommendation in Social Tagging Systems</i>	<i>Cantador, I., Bellogín, A., & Vallet, D. (2010, September). Content-based recommendation in social tagging systems. In Proceedings of the fourth ACM conference on Recommender systems (pp. 237-240).</i>	Jurnal ini mempresentasikan dan mengevaluasi berbagai rekomendasi berbasis konten model yang menggunakan profil pengguna dan item yang ditentukan dalam istilah daftar <i>tag sosial</i> berbobot. Pendekatan yang dipelajari adalah adaptasi informasi <i>Vector Space</i> dan <i>Okapi BM25</i> model pengambilan. Kami membandingkan secara empiris para pemberi rekomendasi menggunakan dua dataset yang diperoleh dari <i>Delicious</i> dan <i>Last.fm</i> sosial sistem, untuk menganalisis kinerja pendekatan di skenario dengan domain berbeda dan perilaku pemberian tag.
2	<i>Content-based News Recommendation</i>	<i>Kompan, M., & Bielíková, M. (2010, September). Content-based news recommendation. In International conference on electronic commerce and web technologies (pp. 61-72). Springer, Berlin, Heidelberg.</i>	Informasi yang berlebihan adalah salah satu masalah terbesar sekarang. Kita bisa melihatnya di berbagai domain, termasuk bisnis khususnya di berita. Ini lebih penting dalam kaitannya dengan web dan berita portal, tempat kualitas portal berita biasanya diukur jumlah berita yang ditambahkan ke situs. Lalu portal berita paling terkenal tambahkan ratusan artikel baru setiap hari. Solusi klasik biasanya digunakan untuk mengatasi overloading informasi adalah rekomendasi. Dalam tulisan ini kami menyajikan pendekatan untuk rekomendasi berita berbasis konten cepat, berbasis pada pencarian <i>cosine similarity</i> .
3	<i>GloVe: Global Vectors for Word Representation</i>	<i>Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language</i>	Metode terbaru untuk mempelajari ruang vektor representasi kata-kata telah berhasil dalam menangkap semantik halus dan keteraturan sintaksis menggunakan aritmatika vektor, tetapi asal dari keteraturan tersebut tetap buram. Kami

		<i>processing</i> (EMNLP) (pp. 1532-1543).	menganalisis dan buat eksplisit properti model yang dibutuhkan agar keteraturan seperti itu muncul dalam kata vektor. Hasilnya adalah model regresi log linear global baru yang digabungkan keunggulan dari dua model utama keluarga dalam literatur: matriks global faktorisasi dan jendela konteks lokal metode.
--	--	--	--

1.6 Sistematika Penulisan

BAB I : PENDAHULUAN

Pada bab ini dijelaskan latar belakang, identifikasi masalah, tujuan, ruang lingkup dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab ini dijelaskan teori-teori dan penelitian terdahulu yang digunakan sebagai acuan dan dasar dalam penelitian.

BAB III : ANALISIS ORGANISASI PERUSAHAAN

Pada bab ini berisi penjelasan tentang profil perusahaan tempat internship yang penulis lakukan.

BAB IV : METODOLOGI PENELITIAN

Pada bab ini berisi tentang alur metodologi yang akan dipakai pada proses pengerjaan internship yang penulis lakukan.

BAB V : ANALISIS DAN PERANCANGAN SISTEM

Pada bab ini berisi analisis dan perancangan sistem dibuat dengan *flowchart*.

BAB IV : PENGKAJIAN DAN EVALUASI

Pada bab ini berisi tentang penjelasan hasil dan evaluasi serta manfaat apa yang didapat dari pengerjaan *internship* ini.

BAB V : PENUTUP

Pada bab ini berisi kesimpulan dan saran perlu disampaikan dari hasil pengerjaan *internship* yang penulis lakukan.

BAB II

LANDASAN TEORI

2.1 *Data Science*

Secara umum, *data science* atau ilmu data adalah penerapan kuantitatif dan metode kualitatif untuk memecahkan masalah yang relevan dan memprediksi hasil. Dengan luas dan semakin banyak data, maka pengetahuan dan analisis data tidak terlepas. Profesor Jeff Stanton dari *Syracuse University* dikutip oleh *Dumbill et al.* (2013): Dari perspektif pengajaran, sebagai anggota fakultas saya bisa ajari seseorang cara melakukan test dalam 10 menit, dan pada data bisa ajari mereka cara menulis program *Python*. *Data science* juga mempelajari bagaimana mengumpulkan data, lalu diolah dengan melakukan *cleaning* dan pengolahan lainnya sesuai dengan tujuan dan tahap akhir adalah dengan membuat visualisasi hasil analisis data tersebut untuk disajikan [5].

2.2 *Python*

Python adalah bahasa pemrograman yang tujuan secara umum bisa digunakan dalam beberapa bidang seperti web, data, statistik, mobile. Sedangkan bahasa *R* umumnya terbatas untuk tugas-tugas dalam bidang komputasi statistik. Dengan *Python*, kita bisa belajar dan membangun keyakinan menggunakan konsep pemrograman umum yang umum di antara bahasa pemrograman populer lainnya seperti *Java*, *C*, atau *C++*. *Python* juga mengintegrasikan pemrograman umum konsep dengan penekanan pada persiapan, pemrosesan, dan presentasi data. Selain itu, *python* juga sebagai pembelajaran yang intensif, seperti pembelajaran mesin atau machine learning. *Python* menjadi suatu alat untuk *data science* atau ilmu data [6].

2.3 *Glove*

Algoritma *GloVe* merupakan salah satu metode representasi vektor kata. Itu membangun *embeddings* kata menggunakan model log bilinear. Algoritma ini

menggabungkan keunggulan suatu text dan faktorisasi matriks global. *GloVe* memiliki banyak *hyperparameter* itu dapat mempengaruhi hasilnya, tetapi ukuran dan dimensi vektor adalah yang paling efektif parameter. [7].

2.4 *Scraping*

Scraping atau *Web Scraping* (Turland, 2010) adalah proses pengambilan sebuah dokumen semi-terstruktur dari internet, umumnya berupa halaman-halaman web dalam bahasa markup seperti *HTML* atau *XHTML*, dan menganalisis dokumen tersebut untuk diambil data tertentu dari halaman tersebut untuk digunakan bagi kepentingan lain [8].

2.5 *Recommendation System*

Recommendation System merupakan sistem yang merekomendasikan suatu hal atau data yang sering kita jumpai sehari-hari, misalnya di *amazon.com* atau *e-commerce* lainnya, kita sering mendapatkan rekomendasi tentang barang yang harus kita beli. Di *youtube* atau *spotify* bisa juga mendapatkan rekomendasi tentang video atau lagu yang harus di putar [9].

2.6 *Pandas*

Pandas adalah alat analisis dan manipulasi data sumber terbuka yang cepat, kuat, fleksibel, dan mudah digunakan, dibangun menggunakan bahasa pemrograman *Python* [10].

2.7 **Rekomendasi Pelatihan**

Rekomendasi pelatihan adalah suatu cara untuk membuat suatu daftar rekomendasi pelatihan, pelatihan disini berarti membuat suatu kegiatan *training* atau pembelajaran kepada sekelompok orang. Dengan adanya rekomendasi pelatihan akan menjadi lebih terarah dan akurat pelatihan yang akan diadakan.

2.8 *Word Embedding*

Word embedding adalah sebuah teknik untuk merepresentasikan arti dari sebuah kata dalam istilah kata lain, seperti yang dicontohkan oleh metode *Word2vec*. *Word embedding* memungkinkan identifikasi kata yang digunakan dalam konteks yang mirip dengan kata yang spesifik. *Word embedding* akan mengubah kata atau text ke dalam bentuk vektor. *Word embedding* juga menjadi sebuah metode yang signifikan di kalangan peneliti dalam pemrosesan bahasa alami (NLP) baru-baru ini [11].

2.9 *Text Matching*

Text matching atau Pencocokan teks adalah salah satu masalah penelitian terpenting di banyak kasus NLP, seperti informasi, jawaban pertanyaan, dan rekomendasi. Di antara berbagai jenis *text matching*, pencocokan teks ke suatu dokumen lainnya [12].

2.10 *Cosine Similarity*

Cosine similarity adalah suatu metode yang diterapkan secara luas di pencarian informasi dan studi terkait. Metode ini memodelkan teks sebagai vektor istilah dan kesamaan antara dua teks berasal nilai cosinus antara vektor istilah dua teks [13].

2.11 *Natural Language Processing*

Natural Language Processing (NLP) adalah bidang penelitian dan aplikasi yang mengeksplorasi bagaimana komputer dapat digunakan untuk memahami dan memanipulasi teks atau ucapan bahasa alami agar berguna sesuatu. Peneliti NLP bertujuan untuk mengumpulkan pengetahuan tentang bagaimana manusia memahami dan menggunakan bahasa sehingga alat dan teknik yang tepat dapat dikembangkan untuk membuat sistem komputer memahami dan memanipulasi bahasa alami untuk melakukan tugas yang diinginkan [14].

BAB III

ANALISIS ORGANISASI PERUSAHAAN

3.1 Sejarah Perusahaan



Gambar 3.1 Logo Telkom

Telekomunikasi Indonesia, Tbk (TELKOM) merupakan BUMN yang bergerak di bidang jasa layanan telekomunikasi dan jaringan di wilayah Indonesia. Sebagai Perusahaan milik negara yang sahamnya diperdagangkan di bursa saham, pemegang saham mayoritas perusahaan adalah Pemerintah Republik Indonesia sedangkan sisanya dikuasai oleh publik. perusahaan penyelenggara bisnis T.I.M.E.S (*Telecommunication, Information, Media, Edutainment and Services*) milik negara yang terbesar di Indonesia, yaitu sebuah portofolio bisnis yang lebih lengkap mengikuti tren perubahan bisnis global di masa datang.

Pada tahun 1882, didirikan sebuah badan usaha swasta penyedia layanan pos dan telegraf. Layanan komunikasi kemudian dikonsolidasikan oleh Pemerintah Hindia Belanda ke jawatan *Post Telegraaf Telefoon* (PTT). Pada tanggal 23 Oktober 1856, dimulai dari layanan jasa telegraf elektromagnetik pertama yang menghubungkan Jakarta (Batavia) dengan Bogor (Buitenzorg). Tahun 2009 momen tersebut dijadikan sebagai patokan hari lahir Telkom.

Pada tanggal 14 November 1995 dilakukan Penawaran Umum Perdana saham Telkom. Sejak itu saham Telkom tercatat dan diperdagangkan di Bursa Efek Jakarta (BEJ) dan Bursa Efek Surabaya (BES) (keduanya sekarang bernama Bursa Efek Indonesia (BEI)), Bursa Saham New York (NYSE) dan Bursa Saham London (LSE). Saham Telkom juga diperdagangkan tanpa pencatatan di Bursa Saham Tokyo. Jumlah saham yang dilepas saat itu adalah 933 juta lembar saham.

Tahun 1999 ditetapkan Undang-undang Nomor 36 Tahun 1999 tentang Telekomunikasi. Sejak tahun 1989, Pemerintah Indonesia melakukan deregulasi di sektor telekomunikasi dengan membuka kompetisi pasar bebas. Dengan demikian, Telkom tidak lagi memonopoli telekomunikasi Indonesia.

Tahun 2001 Telkom membeli 35% saham Telkomsel dari PT Indosat sebagai bagian dari implementasi restrukturisasi industri jasa telekomunikasi di Indonesia yang ditandai dengan penghapusan kepemilikan bersama dan kepemilikan silang antara Telkom dan Indosat. Sejak bulan Agustus 2002 terjadi duopoli penyelenggaraan telekomunikasi lokal. Pada 23 Oktober 2009, Telkom meluncurkan “New Telkom” (“Telkom baru”) yang ditandai dengan penggantian identitas perusahaan [15].

3.2 Visi dan Misi Perusahaan

3.2.1 Visi

“Menjadi digital telco pilihan utama untuk memajukan masyarakat.

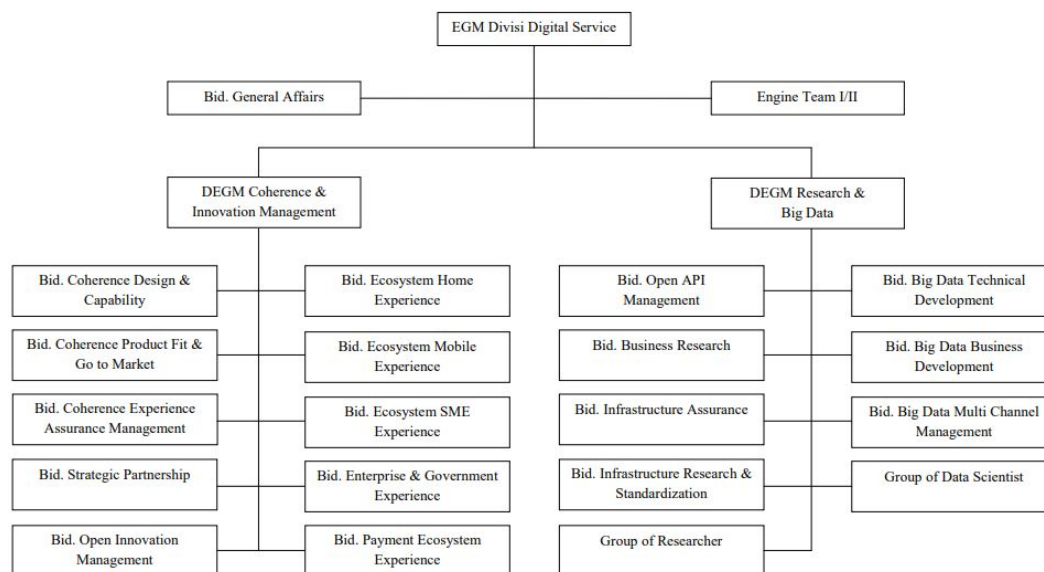
3.2.2 Misi

- Mempercepat pembangunan Infrastruktur dan platform digital cerdas yang berkelanjutan, ekonomis, dan dapat diakses oleh seluruh masyarakat.
- Mengembangkan talenta digital unggulan yang membantu mendorong kemampuan digital dan tingkat adopsi digital bangsa.
- Mengorkestrasi ekosistem digital untuk memberikan pengalaman digital pelanggan terbaik [15].

3.3 Strategi Perusahaan

Dalam upaya bertransformasi menjadi *digital telecommunication company*, *Telkom Group* mengimplementasikan strategi bisnis dan operasional perusahaan yang berorientasi kepada pelanggan (*customer-oriented*). Transformasi tersebut akan membuat organisasi *Telkom Group* menjadi lebih *lean* (ramping) dan *agile* (lincah) dalam beradaptasi dengan perubahan industri telekomunikasi yang berlangsung sangat cepat. Organisasi yang baru juga diharapkan dapat meningkatkan efisiensi dan efektivitas dalam menciptakan *customer experience* yang berkualitas [15].

3.4 Struktur Organisasi



Gambar 3.2 Struktur Organisasi DDB Telkom

3.4.1 Job Description Peserta Internship

Pada kegiatan internship ini penulis melakukan program internship I di PT. Telekomunikasi Indonesia dan ditempatkan pada *Group of Data Scientists*, yang sebagaimana tugas divisi tersebut sebagai divisi mengolah data dan menciptakan suatu solusi dari berbagai masalah yang ada pada pekerjaan yang dilakukan oleh PT. Telkom.

3.5 Ruang Lingkup Internship I

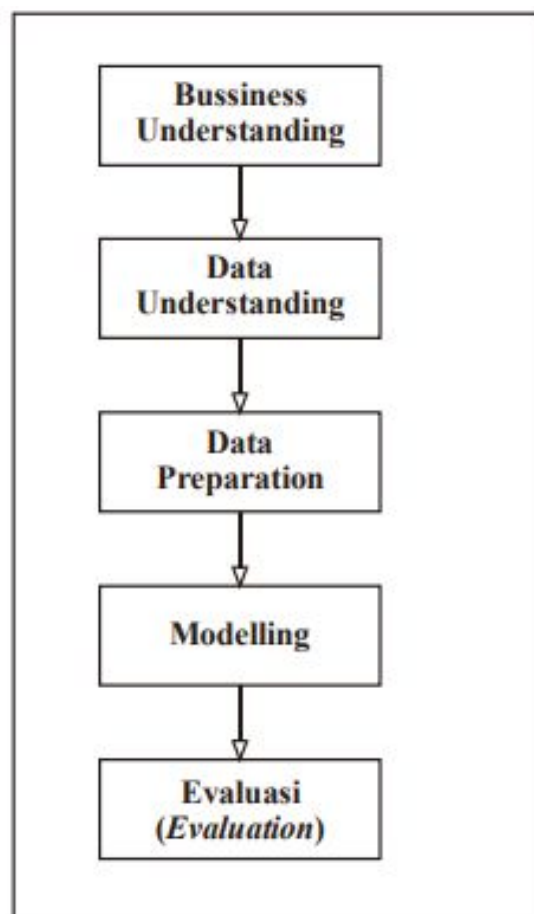
Pada program internship ini penulis berada dalam divisi *Data Scientists* di PT. Telekomunikasi Indonesia. Dimana tugasnya membantu sebuah *project* perusahaan untuk menganalisis *Learning Analytic Telkom Corporate University*. Disini, penulis menjadi bagian dari pengolahan data untuk menerapkan *recommendation system* dalam merekomendasikan pelatihan apa yang cocok oleh *Telkom Corporate University*. Penulis akan mengumpulkan, mengolah dan memvisualisasikan data analisis.

BAB IV

METODOLOGI PENELITIAN

4.1. Diagram Alur Metodologi Penelitian

Metode penelitian yang proses pengerjaan internship I penulis mengambil referensi dengan mengikuti tahapan model *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. Adapun tahapan *CRISP-DM*, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment* (CRISP, 2009). Dalam internship ini penulis membuat beberapa tahapan, sebagai berikut :



Gambar 4.1 Diagram alur penelitian

4.2. Tahapan – Tahapan Diagram Alur Metodologi Penelitian

4.2.1. *Business Understanding*

Business understanding adalah pemahaman terhadap tujuan bisnis, penilaian situasi dan menerjemahkan tujuan bisnis ke dalam tujuan *data mining*. Dalam internship ini membutuhkan pengetahuan untuk mendapatkan data dari *course-course* yang diadakan secara online, *keyword* dari paper dan topik dari setiap referensi *course*. Tidak lupa pula melakukan analisa untuk mendukung proses rekomendasi pelatihan menggunakan dataset *Telkom Corporate University*.

4.2.2. *Data Understanding*

Pada tahap ini dilakukan proses pengumpulan data, kemudian melakukan analisa data serta melakukan evaluasi kualitas data yang digunakan dalam pengerjaan internship. Sumber data yang digunakan dalam internship ini merupakan data-data yang terdapat dari berbagai media. Maka sumber literatur banyak didapatkan dari buku, paper atau jurnal, karya ilmiah dan situs-situs penunjang.

4.2.3. *Data Preparation*

Data Preparation atau bisa disebut juga dengan data preprocessing adalah suatu proses/langkah yang dilakukan untuk membuat data mentah menjadi data yang berkualitas (input yang baik untuk *data mining tools*). Dengan seperti ini proses pengerjaan internship akan lebih mudah.

4.2.4 *Modelling*

Dalam tahap ini, berbagai macam metode pemodelan dipilih dan diterapkan ke dataset yang sudah disiapkan untuk mengatasi kebutuhan bisnis yang sesuai. Adapun metode yang digunakan yaitu :

- *Scraping data*
- *Cleaning data*
- *Word embedding*
- *Extraction keyword*

- *Recommendation system*
- *Content-based filtering*

4.2.5. Evaluation

Pada tahap terakhir ini, model yang sudah dibuat diuji dan dievaluasi keakuratan. Tahap ini mengukur sejauh mana model yang sudah dipilih memenuhi sasaran-sasaran bisnis dan bila demikian, sejauh manakah itu (apakah perlu lebih banyak model untuk dibuat dan diukur).

BAB V

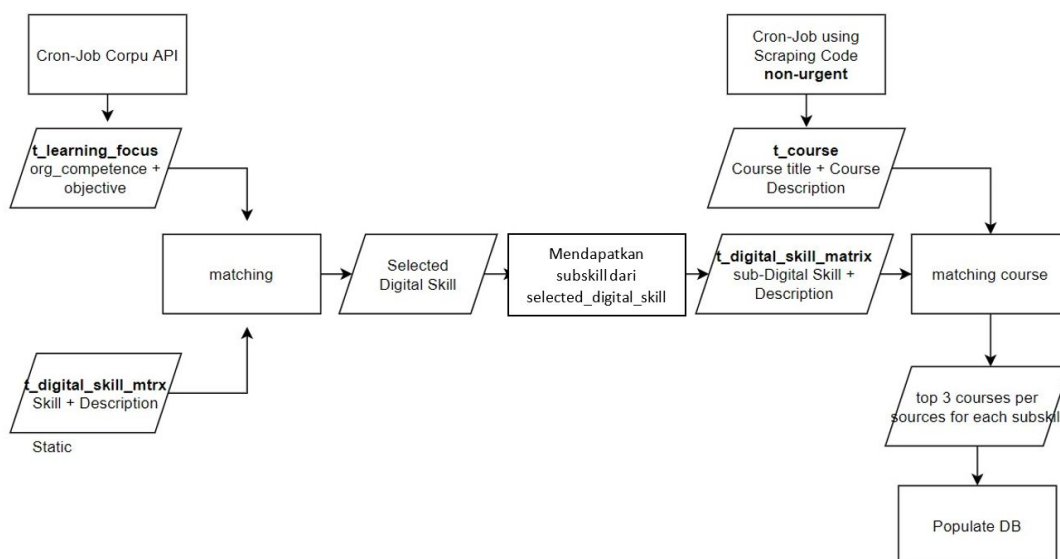
ANALISIS DAN PERANCANGAN SISTEM

5.1. Analisis Sistem yang akan dibangun

Analisis sistem yang akan dibangun dimaksud disini berupa analisis *Flowchart* mengenai sistem yang akan dibangun meliputi proses *recommendation system* untuk mendapatkan hasil rekomendasi pelatihan atau *training recommendation* yang akan diadakan di Telkom Corporate University Bandung serta rekomendasi pembicara atau *expert recommendation*. Adapun *flowchart* yang akan dibangun adalah sebagai berikut:

5.1.1. Analisis Prosedur yang akan dibangun

5.1.1.1 Analisis *Flowchart* pada proses *Training Recommendation*



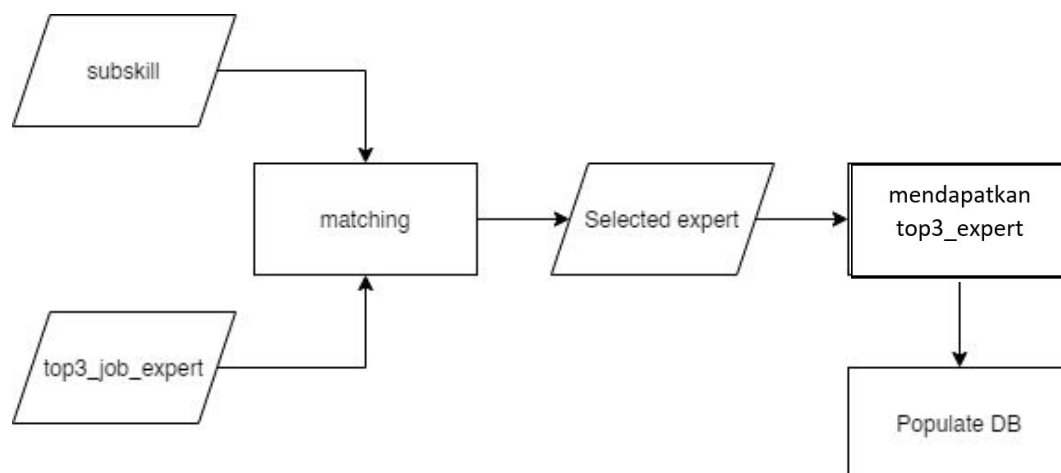
Gambar 5.1 *Flowchart Training Recommendation*

Pada proses *training recommendation* menggunakan beberapa sumber data yaitu database Telkom Corporate University Bandung dan hasil *scraping dataset course online*. Mekanisme kerja yang dibuat adalah sebagai berikut :

1. *t_learning_focus* di *matching* dengan *t_digital_skill_mtrx*, melibatkan data *org_competence+objective* dan *skill+description*, disini dilakukan *matching* dengan mengambil kemiripan antara kedua data tersebut.

2. Jika *matching* berhasil maka akan mendapatkan *digital skill* yang *make sense* dengan kedua data yang di *matching* sebelumnya.
3. Kemudian mengambil *subskill* dari *digital skill*.
4. Melakukan *matching* dengan data *subskill+description* dengan *course_title + course_description*.
5. Jika *matching* berhasil maka menghasilkan sebuah *training recommendation* yang berisikan *course online* yang *make sense* dengan data yang di *matching*.
6. Terakhir data akan dimasukkan ke dalam *database*.

5.1.1.2 Analisis *Flowchart* pada proses *Expert Recommendation*



Gambar 5.2 *Flowchart Expert Recommendation*

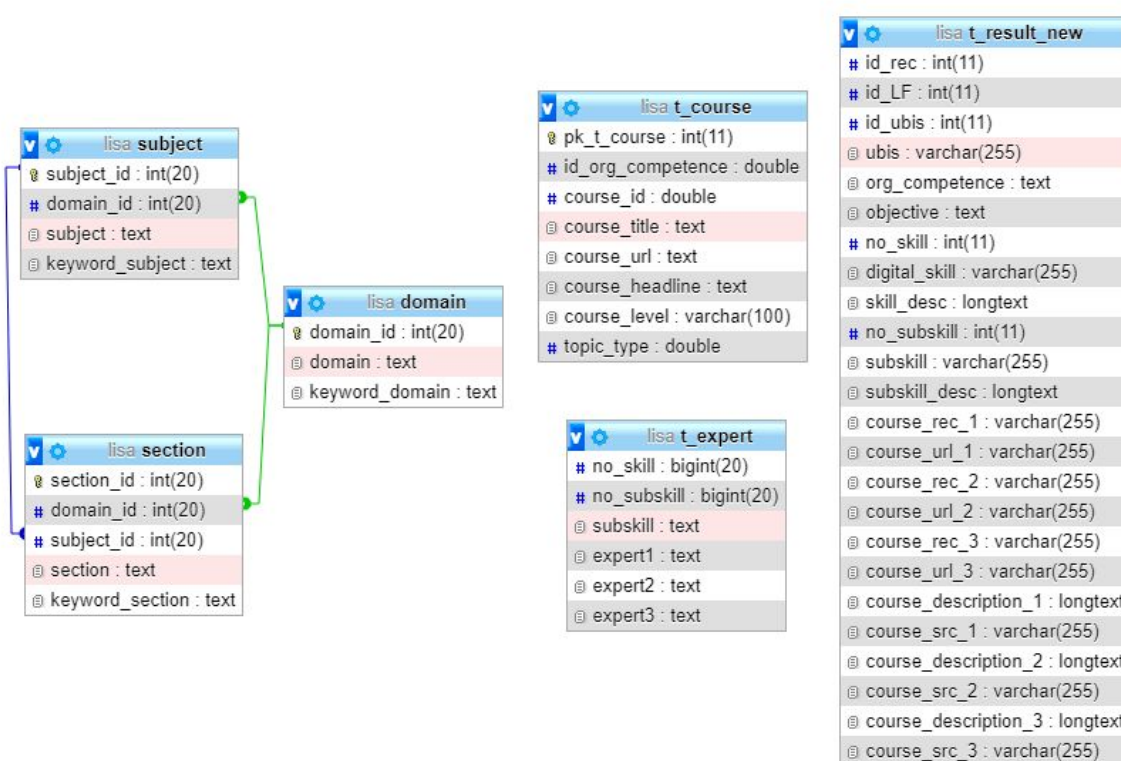
Pada proses *expert recommendation* menggunakan beberapa sumber data yaitu *database Telkom Corporate University Bandung* dan hasil *scraping dataset expert LinkedIn*. Mekanisme kerja yang dibuat adalah sebagai berikut :

1. Melakukan *matching subskill* dengan data *top3_job* setiap *expert* yang sudah di *scraping*.
2. Jika berhasil maka akan mendapatkan *expert* yang *make sense* dengan data yang di *matching*.
3. Dapatin *top3 expert* yang data *job* yang dimiliki lebih mempunyai kemiripan dengan *subkill* yang di *matching*.

4. Terakhir data akan dimasukkan ke dalam *database*.

5.2. Perancangan Basis Data/*Database* Sistem.

Perlu diketahui bahwa jumlah *database* untuk keseluruhan ada 80 *table* tetapi penulis akan memperlihatkan beberapa rancangan *database* untuk sistem yang diperlukan *Learning Data Analytics Dashboard* ialah sebagai berikut :



Gambar 5.3 Perancangan *Database*

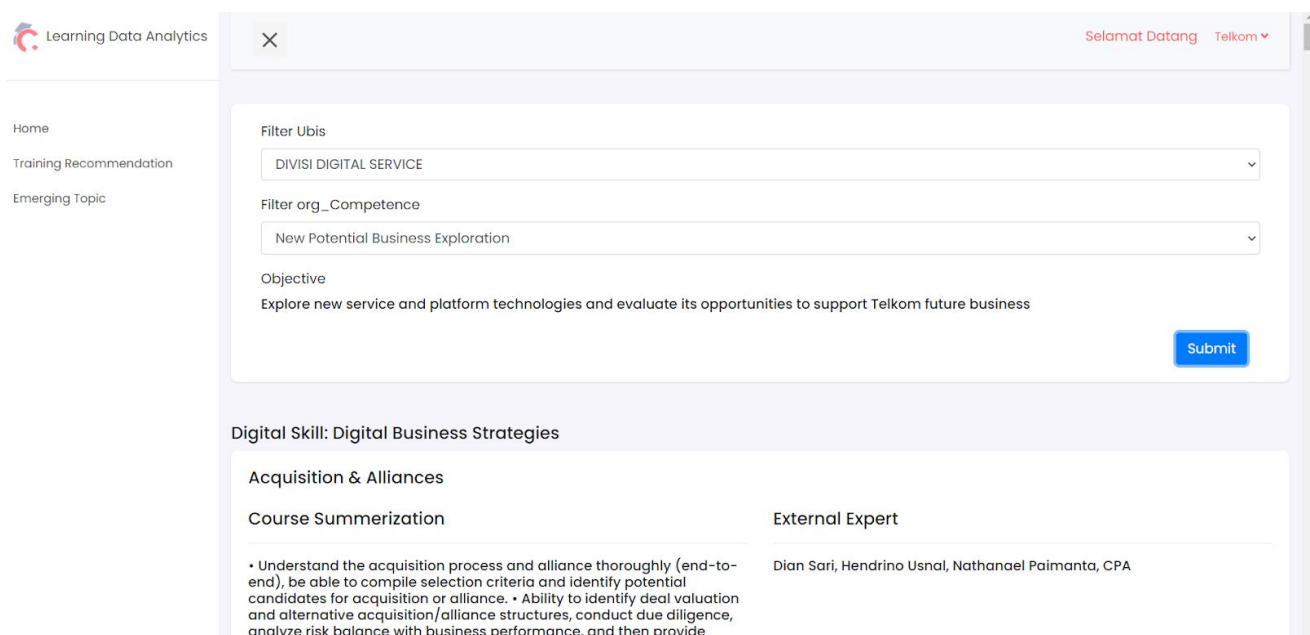
5.3. Perancangan *User Interface* Sistem.

5.3.1 Perancangan Tampilan *Home*



Gambar 5.4 Perancangan Tampilan *Home*

5.3.2 Perancangan Tampilan *Menu Training Recommendation*



Gambar 5.5 Perancangan Tampilan *Menu Training Recommendation*

5.3.3 Perancangan Tampilan Menu Emerging Topic

Learning Data Analytics

Selamat Datang Telkom

Home

Training Recommendation

Emerging Topic

Top 6 Topic Recommendation

Select Domain Select Subject Select Section Submit

No	Keyword
1	gait
2	minimally invasive surgery
3	assistive devices
4	exoskeleton
5	variable stiffness
6	rehabilitation robotics

©2020 All Rights Reserved

Main Office
Lorem ipsum dolor sit amet, consectetur adipiscing elit.
Nunc mattis diam vel ipsum viverra mattis.

Bandung Office
Lorem ipsum dolor sit amet, consectetur adipiscing elit.
Nunc mattis diam vel ipsum viverra mattis.

Gambar 5.6 Perancangan Tampilan *Menu Emerging Topic*

5.4. Perancangan Arsitektur Perangkat Lunak dan Perangkat Keras Sistem.

Analisis kebutuhan perangkat lunak dan perangkat keras dilakukan untuk mengetahui spesifikasi kebutuhan yang diperlukan sistem. Spesifikasi kebutuhan melibatkan analisis perangkat keras (*hardware*) dan analisis perangkat lunak (*software*). Adapun kebutuhan yang akan dibuat adalah sebagai berikut :

5.4.1 Kebutuhan Perangkat Lunak

Adapun spesifikasi perangkat lunak (*software*) yang dilakukan saat proses eksperimen untuk mendukung sistem yang akan dibangun adalah sebagai berikut :

Tabel 5.1: Kebutuhan Perangkat Lunak

No	Jenis Perangkat Lunak	Spesifikasi
1	Sistem Operasi	<i>Windows 10</i>
2	Server	<i>Google Colab (Python)</i>

3	Bahasa Pemrograman	<i>Python, PHP</i>
4	<i>Browser</i>	<i>Chrome</i>
5	<i>Editor Text</i>	<i>Google Colab, Sublime Text, Overleaf Latex</i>
6	<i>Database</i>	<i>Mysql</i>

5.4.2 Kebutuhan Perangkat Keras

Adapun spesifikasi perangkat keras (hardware) yang dilakukan saat proses eksperimen untuk mengimplementasikan sistem yang akan dibangun adalah sebagai berikut:

Tabel 5.2: Kebutuhan Perangkat Keras

No	Jenis Perangkat Keras	Spesifikasi
1	<i>Harddisk</i>	<i>256 GB</i>
2	<i>RAM</i>	<i>8 GB</i>
3	<i>Processor</i>	<i>Core i7</i>
4	<i>VGA</i>	<i>2 GB</i>
5	<i>Python</i>	<i>Google Colab</i>
6	<i>Mouse dan Keyboard</i>	<i>Standart</i>

BAB VI

PENGKAJIAN DAN EVALUASI

Langkah pengkajian dan evaluasi merupakan hal terpenting yang harus dilaksanakan guna mendapatkan hasil yang maksimal dari sistem yang sudah dibangun. Penulis akan memperlihatkan hasil secara keseluruhan serta evaluasi terkait sistem yang sudah dibuat. Berdasarkan perancangan sistem yang disusun, penulis menggunakan *Google Colab* sebagai alat untuk eksperimen dengan menggunakan bahasa pemrograman *Python*. Setelah dari *google colab* hasil akan dimasukkan ke *database* serta ditampilkan ke *web*.

6.1 Hasil Secara Keseluruhan

Filter Ubis

DIVISI DIGITAL SERVICE

Filter org_Competence

New Potential Business Exploration

Objective

Explore new service and platform technologies and evaluate its opportunities to support Telkom future business

Submit

Gambar 6.1 *Filter* untuk *Recommendation System*

Digital Skill: Digital Business Strategies	
Acquisition & Alliances	
Course Summerization	External Expert
<ul style="list-style-type: none"> • Understand the acquisition process and alliance thoroughly (end-to-end), be able to compile selection criteria and identify potential candidates for acquisition or alliance. • Ability to identify deal valuation and alternative acquisition/alliance structures, conduct due diligence, analyze risk balance with business performance, and then provide related recommendations to the company. • The ability to analyze a digital product/service is developed or through Acquisition & Alliances. 	Dian Sari, Hendrino Usnal, Nathanael Paimanta, CPA

Gambar 6.2 Hasil *Expert Recommendation*

Growth and Acquisition Strategy		
Course Description	Course Url	Source Course
School of Business. Get started with the first course in the Growth Product Manager Nanodegree program. Hone specialized Product Management skills in growth and acquisition strategy by learning how to build an agile acquisition plan with market-based measurable KPIs which fits well into the overall growth strategy. Apply the fundamentals of acquisition funnels and growth loops, measuring and optimizing acquisition performance, defining target audiences and key acquisition channels, and A/B testing at scale to expand the user base of your product.	https://www.udacity.com/course/growth-and-acquisition-strategy--nd037-1	Udacity

Gambar 6.3 Hasil *Training Recommendation*

6.2 Kemanfaatan fungsi sistem terhadap permasalahan

Adapun manfaat dari sistem yang dibuat terhadap permasalahan ialah :

1. *Learning Data Analytics Dashboard* dapat menjadi suatu solusi untuk Telkom Corporate University Bandung dalam menganalisis pelatihan apa yang cocok diadakan di lingkungan Telkom.
2. *Training & Expert Recommendation* dapat menjadi saran topik dan tema pelatihan yang akan diadakan.
3. *Emerging Topic Recommendation* dapat menjadi saran topik yang lagi *trending*.

BAB VII

PENUTUP

7.1 Kesimpulan dan Saran

7.1.1 Kesimpulan

Berdasarkan hasil analisis pada laporan internship I ini, maka penulis membuat beberapa kesimpulan sebagai berikut :

1. Metode *recommendation system* dapat menjadi solusi dalam merekomendasikan daftar pelatihan yang akan dibuat.
2. Metode *recommendation system* dapat mengolah dan menganalisis data dengan baik.
3. *Learning Data Analytics Dashboard* dapat menjadi suatu solusi untuk Telkom Corporate University Bandung dalam menganalisis pelatihan apa yang cocok diadakan di lingkungan Telkom.

7.2.2 Saran

Berdasarkan hasil Kesimpulan diatas, maka penulis membuat beberapa saran untuk diterapkan di penelitian berikutnya, sebagai berikut :

1. Metode *recommendation system* bisa lebih ditingkatkan lagi keakuratannya dalam *word embedding* dan *cosine similarity* dengan menggunakan model *GloVe*.
2. Metode *recommendation system* diharapkan di penelitian berikutnya dapat menghasilkan suatu olahan data yang lebih baik.
3. *Learning Data Analytics Dashboard* kedepannya diharapkan bisa mempunyai *User Interface* yang lebih *responsive*.

7.2 Lampiran-lampiran

7.2.1 Laporan Kemajuan Bimbingan (Kambing)



POLITEKNIK POS INDONESIA
PROGRAM STUDI D4 TEKNIK INFORMATIKA
JL. SARIASIH NO. 54 BANDUNG - 40151
Telp. 022-2009562, 2009570
Fax. 022-2009568

FORMULIR KEGIATAN

INTERNSHIP I

TA. 2020/2021

Nama : Muhammad Fahmi
Npm : 1174021
Judul : Cosine Similarity With Word Embedding For Learning
Data Analytics Dashboard Using GloVe
Pembimbing : Rolly Maulana Awangga S.T., M.T.



Pertemuan	Tanggal	Sudah Dikerjakan	Pekerjaan Selanjutnya	Nilai
1	20-10-2020	diskusi tema dan judul	consine similarity glove ppji chapter1	100
2	30-10-2020	progress ppji chapter 1	explore jurnal dan ppji 4 1	100
3	07-11-2020	progress ppji 4 1 related works	melanjutkan ppji 4 1 fokus ke grammatly beli akun grammatly	100
4	14-11-2020	progress ppji 4 2 introduction	melanjutkan ppji 4 3	100
5	-	-	-	-
6	28-11-2020	progress ppji 4 3 methods	melanjutkan ppji 4 4 experiments	100
7	-	-	-	-
8	08-12-2020	progress ppji 4 4 experiments	melanjutkan evaluasi 4 4 experiments dan cari referensi judul	100
9	18-12-2020	progress ppji 4 5 result	melanjutkan evaluasi 4 5 result dan cari mencari referensi jurnal	100
10	24-12-2020	progress jurnal referensi	melanjutkan evaluasi referensi jurnal	100
11	30-12-2020	progress jurnal finishing	melanjutkan evaluasi jurnal explore dan ubah template ke elsevie	100
12	07-01-2021	finishing jurnal elsevier	melanjutkan evaluasi jurnal	100
			Rata-Rata:	83.33

Bandung, 07 Januari 2021

Pembimbing,



Rolly Maulana Awangga S.T., M.T.
NIDN. 0410118609

7.2.2 Penilaian Dari Pembimbing Eksternal

Lampiran 14 : FORMULIR PENILAIAN DARI PEMBIMBING EKSTERNAL


FORMAT PENILAIAN INTERNSHIP PROGRAM STUDI D4 TEKNIK INFORMATIKA POLITEKNIK POS INDONESIA				
NAMA		NPM	Tempat Tgl. Lahir	
MUHAMMAD FAHMI		1174021	Kisaran, 17 Juni 1999	
JUDUL INTERNSHIP		: <i>Cosine Similarity With Word Embedding For Learning Data Analytics Dashboard Using Glove</i>		
PEMBIMBING EKSTERNAL		: Faisal Baharuddin		
NO.	KOMPONEN YANG DINILAI	NILAI MAKS	PENILAIAN (ANGKA)	RATA-RATA
1.	PENAMPILAN INDIVIDUAL			6
	A. Penampilan Berpakaian	7	6	
	B. Sikap Terhadap Orang Lain	8	7	
	C. Semangat Bekerja	7	7	
	D. Kematangan Dalam Bertindak	6	5	
	E. Adaptasi Tempat Kerja	6	5	
	F. Pengetahuan Yang Mendukung Pekerjaan	6	5	
	G. Kehadiran Ditempat Kerja	8	7	
2.	KINERJA PKL			6.8
	A. Ketelitian & Ketepatan Dalam Pekerjaan	8	7	
	B. Kualitas Produk / Kerja	8	7	
	C. Kemandirian Dalam Melaksanakan Pekerjaan	7	7	
	D. Kemampuan Bekerjasama	7	7	
	E. Kemampuan Berkomunikasi	8	8	
	F. Manajemen Waktu	7	6	
	G. Kemampuan Menganalisa Masalah	7	6	
TOTAL		100	90	

BANDUNG, 06 JANUARI 2020
PEMBIMBING EKSTERNAL



FAISAL BAHARUDDIN.
NIK. 651196

7.2.3 Curriculum Vitae dalam bahasa Inggris dan Indonesia



MUHAMMAD FAHMI

for more details : <https://fahmifc123.github.io>

PROFILE

I was born in Kisaran, Asahan, North Sumatra and now I am still a student at Politeknik Pos Indonesia Bandung. I am pursuing the fields of Data Science, AI and Machine Learning. All because it is in accordance with the programming language that I am currently working on, namely Python. Apart from that, there are many things I want to learn more, especially in the field of data.

EDUCATION

FORMAL :

2017 - NOW D4 INFORMATICS ENGINEERING (IPK 3,50)

2014 - 2017 SMAN 2 KISARAN (IPA)

2011 - 2014 SMPN 6 KISARAN

2005 - 2011 SD MIN KISARAN

NON FORMAL :

2020 - DATA SCIENCE COURSE DQLAB.ID

CERTIFICATION :

2019 - 2022 COMPETENCY CERTIFICATE JUNIOR WEB PROGRAMMER BNSP

INTERNSHIP :

2020-2021 DATA SCIENTIST DDB TELKOM

PENGALAMAN

ORGANIZATION :

2018 - 2019 HEAD OF DEPARTMENT KOMINFO BEM

2018 - 2019 HEAD OF DEPARTMENT SYI'AR LDK

2015 - 2016 HEAD OF ROHIS SMAN 2 KISARAN

PROJECT :

PROYEK I IMPLEMENTATION SMART BOX LOGIST K

PROYEK II HUMAN BRAIN WAVE DETECTION


PROYEK III SMART DOOR FACE RECOGNITION


VOLUNTEERS AND TRAINING :


2019 SPEAKER OF PYTHON TRAINING


2018 COMMUNITY DEDICATION IYALE


KONTAK

 Muhammad Fahmi

 +62 812-6329-9950

 mfahmipamungkas123@gmail.com

 Muhammad Fahmi Pamungkas

 @mfahmipamungkas




BASIC SKILLS

TEAM WORK	██████████
CREATIVITY	██████████
PROBLEM SOLVING	██████████
PUBLIC SPEAKING	██████████
CRITICAL	██████████

TECHNICAL SKILLS

PYTHON	██████████
R	██████████
PHP	██████████
PANDAS	██████████
NUMPY	██████████
OPENCV	██████████
MATPLOTLIB	██████████
MS OFFICE	██████████

HOBBIES

DAFTAR PUSTAKA

- [1] “Pengertian Data, Fungsi Data dan Macam Jenisnya”. pendidikan.co.id. 20 Juli 2020. <https://pendidikan.co.id/pengertian-data/>
- [2] Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big data*, 1(1), 51-59.
- [3] Yang, S., Korayem, M., AlJadda, K., Grainger, T., & Natarajan, S. (2017). Combining content-based and collaborative filtering for job recommendation system: A cost-sensitive Statistical Relational Learning approach. *Knowledge-Based Systems*, 136, 37-45.
- [4] “Recommendation System Dengan Python : Definisi (Part 1)”. medium.com. 11 Juli 2019. <https://medium.com/data-folks-indonesia/recommendation-system-dengan-python-definisi-part-1-71154dc3f700>
- [5] Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big data*, 1(1), 51-59.
- [6] Brunner, R. J., & Kim, E. J. (2016). Teaching data science. *Procedia Computer Science*, 80, 1947-1956.
- [7] Ibrahim, M., Gauch, S., Salman, O., & Alqahatani, M. (2020). Enriching Consumer Health Vocabulary Using Enhanced GloVe Word Embedding. *arXiv preprint arXiv:2004.00150*.
- [8] Josi, A., & Abdillah, L. A. (2014). Penerapan teknik web scraping pada mesin pencari artikel ilmiah. *arXiv preprint arXiv:1410.5777*.
- [9] “Sekilas Tentang Recommender System” . <https://structilmy.com/>. 25 Juni 2019. <https://structilmy.com/2019/06/sekilas-tentang-recommender-system/>
- [10] “Pandas”. <https://pandas.pydata.org/>.
- [11] Ganguly, D., Roy, D., Mitra, M., & Jones, G. J. (2015, August). Word embedding based generalized language model for information retrieval. In *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval* (pp. 795-798).

- [12] Jiang, J. Y., Zhang, M., Li, C., Bendersky, M., Golbandi, N., & Najork, M. (2019, May). Semantic text matching for long-form documents. In The World Wide Web Conference (pp. 795-806).
- [13] Rahutomo, F., Kitasuka, T., & Aritsugi, M. (2012, October). Semantic cosine similarity. In The 7th International Student Conference on Advanced Science and Technology ICAST (Vol. 4, No. 1)
- [14] Chowdhary, K. R. (2020). Natural language processing. In Fundamentals of Artificial Intelligence (pp. 603-649). Springer, New Delhi.
- [15] “Tentang Telkomgroup” . telkom.co.id. https://www.telkom.co.id/sites/about-telkom/id_ID/page/profil-dan-riwayat-singkat