

**PENGEMBANGAN MODEL *MACHINE LEARNING* BERBASIS
UNIVERSAL SENTENCE ENCODER UNTUK MENINGKATKAN
AKURASI ALGORITMA PENCOCOKAN PADA *PLATFORM*
SOSIAL-KOLABORATIF COLLABOLIO**

SKRIPSI

Disusun sebagai bentuk pelaporan penelitian skripsi dan salah satu syarat
menempuh Sarjana Strata 1 (S1)



Disusun oleh:

AMELIA NUR SAFITRI

NPM. 3332200010

**PROGRAM STUDI TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS SULTAN AGENG TIRTAYASA**

2024

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya sebaga penulis skripsi berikut:

Judul : Pengembangan Model *Machine Learning* Berbasis
Universal Sentence Encoder Untuk Meningkatkan Akurasi
Algoritma Pencocokan Pada Platform Sosial-kolaboratif
Collabolio

Nama Mahasiswa : Amelia Nur Safitri

NPM : 3332200010

Fakultas/Jurusan : Teknik/Teknik Elektro

Menyatakan dengan sesungguhnya bahwa skripsi di atas adalah benar-benar hasil karya asli saya dan tidak memuat hasil karya orang lain, kecuali dinyatakan melalui rujukan yang benar dan dapat dipertanggungjawabkan. Apabila dikemudian hari ditemukan hal-hal menunjukkan bahwa sebagian atau seluruh karya ini bukan karya saya, maka saya bersedia dituntut melalui hukum yang berlaku. Saya juga bersedia menanggung segala akibat hukum yang timbul dari pernyataan yang secara sadar dan sengaja saya nyatakan melalui lembar ini.

Cilegon. 4 Juni 2024



Amelia Nur Safitri
NPM. 3332200010

LEMBAR PENGESAHAN

Dengan ini ditetapkan bahwa skripsi berikut:

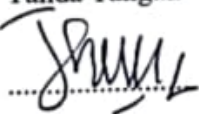
Judul : Pengembangan Model *Machine Learning* Berbasis
Universal Sentence Encoder Untuk Meningkatkan Akurasi
Algoritma Pencocokan Pada Platform Sosial-kolaboratif
Collabolio


Nama Mahasiswa : Amelia Nur Safitri

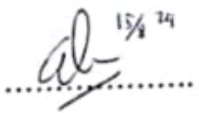
NPM : 3332200010

Fakultas/Jurusan : Teknik/Teknik Elektro

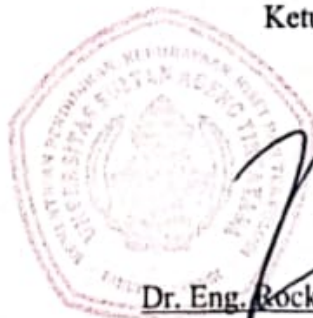
Telah diuji dan dipertahankan pada tanggal 25 Juli 2024 melalui Sidang Skripsi di
Fakultas Teknik Universitas Sultan Ageng Tirtayasa Cilegon dan dinyatakan
LULUS/HIDAK LULUS:

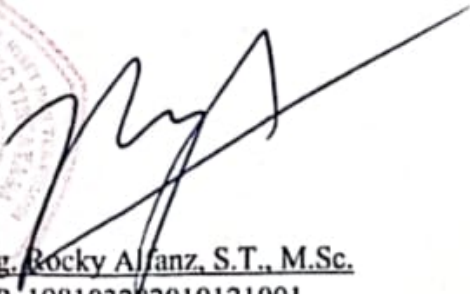
Pembimbing I : Dr. Eng. Teguh Firmansyah, S.T., M.T. ^{23/7} Tanda Tangan 

Penguji I : Masjudin, S.T., M. Eng. 

Penguji II : Ceri Ahendyarti, S.T., M. Eng. ^{15/7} 

Mengetahui,
Ketua Jurusan




Dr. Eng. Rocky Alfanz, S.T., M.Sc.
NIP. 198103262010121001

PRAKATA

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan penelitian. Penelitian ini dilakukan dalam rangka Kegiatan Merdeka Belajar Kampus Merdeka pada tingkat Srata 1 (S1) di Universitas Sultan Ageng Tirtayasa.

Saya menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak lain. Oleh karena itu, penulis ingin mengucapkan terimakasih sebanyak-banyaknya kepada semua pihak yang telah membantu dalam pelaksanaan, penulisan laporan, dan penyelesaian penelitian ini, yaitu:

1. Allah SWT atas Segala nikmat dan rahmat serta karunia pertolongan-Nya selama penulisan skripsi ini.
2. Dr. Eng. Rocky Alfanz, S.T., M.Sc., selaku Ketua Jurusan Teknik Elektro Fakultas Teknik Universitas Sultan Ageng Tirtayasa.
3. Dr. Suhendar, S.Pd., M.T., selaku Dosen Pembimbing Akademik penulis.
4. Dr. Eng. Teguh Firmansyah, S.T., M.T. IPM., selaku Dosen Pembimbing Skripsi yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penulis.
5. Deti Anggraeni Ekawati, selaku Pembimbing Lapangan dari Bangkit Academy yang telah memberikan bimbingan penulis selama kegiatan penelitian.
6. Cinta pertama dan panutanku, Ayahanda Asmuni dan pintu surgaku Ibunda Harni. Terimakasih atas segala pengorbanan dan tulus kasih yang diberikan. Beliau memang tidak sempat merasakan pendidikan bangku perkuliahan, namun mereka mampu senantiasa memberikan yang terbaik, tak kenal lelah mendoakan serta memberikan perhatian dan dukungan sehingga penulis mampu menyelesaikan studinya sampai meraih gelar sarjana. Semoga ayah dan mama sehat, panjang umur, dan bahagia selalu.
7. Kepada cinta dan kasih kedua saudara kandung saya, Najwa Fatihah Sari dan Ahmad Prabowo Putra Hanif. Terimakasih atas segala doa, usaha, dan motivasi yang telah diberikan kepada kakak perempuanmu ini.

8. Rekan-rekan mahasiswa Program Studi Teknik Elektro 2020 atas dukungan dan kerjasama dalam menempuh pendidikan serta penyelesaian skripsi ini.

Penulis sangat menyadari bahwa hasil dari penelitian ini masih memiliki kekurangan, namun penulis berharap hasil dari penelitian ini dapat bermanfaat bagi pembaca, khususnya bagi penulis sendiri. Oleh karena itu berbagai kritik dan saran, sangat penulis harapkan untuk kemajuan hasil penelitian ini.

Cilegon, 22 Maret 2024

Amelia Nur Safitri

ABSTRAK

Amelia Nur Safitri

3332200010

Teknik Elektro

Pengembangan Model *Machine Learning* Berbasis *Universal Sentence Encoder* untuk Meningkatkan Akurasi Algoritma Pencocokan Pada Platform Sosial-Kolaboratif Collabolio

Pandemi COVID-19 meningkatkan pengangguran di Indonesia sejak tahun 2020. Tingkat pengangguran terbuka mencapai 7,07% pada Februari 2021. Penggunaan teknologi berkontribusi pada tingginya pengangguran di kalangan lulusan perguruan tinggi yang kurang memiliki *soft skill*. Collabolio menggunakan *Universal Sentence Encoder* (USE) dan *Tensorflow Recommenders* (TFRS) untuk mencocokkan kemampuan dan minat pekerjaan pengguna, sehingga memungkinkan kolaborasi antara individu dengan ide proyek. *Universal Sentence Encoder* menghasilkan skor kemiripan antara *item* dalam dataset, sedangkan *TensorFlow Recommenders* digunakan untuk membangun model sistem rekomendasi yang menangani masalah *cold-start* dan pembelajaran yang tidak terawasi. Penggunaan USE dan TFRS dalam Collabolio menghasilkan prediksi yang baik dengan total *loss* sebesar 5.50.

Kata Kunci : USE, TFRS, Sistem Rekomendasi

ABSTRACT

Amelia Nur Safitri

3332200010

Electrical Engineering

Development of a machine learning model based on a universal sentence encoder to improve the accuracy of the matching algorithm on the social collaboration platform Collabolio.

The COVID-19 pandemic has increased unemployment in Indonesia since 2020. The open unemployment rate reached 7.07% in February 2021. The use of technology contributes to high unemployment among graduates who lack soft skills. Collabolio uses Universal Sentence Encoder (USE) and *Tensorflow* Recommenders (TFRS) to match users' job skills and interests, enabling collaboration between individuals and project ideas. The Universal Sentence Encoder generates similarity scores between items in the dataset, while *Tensorflow* Recommenders is used to build a recommendation system model that addresses cold-start and unsupervised learning issues. Using USE and TFRS in Collabolio results in a good prediction with an overall loss of 5.50.

Keywords: USE, TFRS, Recommendation System

DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PERNYATAAN KEASLIAN SKRIPSI	ii
LEMBAR PENGESAHAN	iii
PRAKATA.....	iv
ABSTRAK	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	ix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	5
1.6 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 <i>Artificial Intelligence</i>	7
2.2 <i>Machine Learning</i>	8
2.3 Pengolahan data.....	10
2.4 Bahasa Pemrograman Python.....	12
2.5 <i>Deep Learning</i>	12
2.6 <i>Natural Language Processing (NLP)</i>	13
2.7 <i>Word Embedding</i>	15
2.8 <i>Collaborative Filtering</i>	18
2.9 <i>Universal Sentence Encoder (USE)</i>	19
2.10 <i>Tensorflow Recommendation</i>	20
2.11 Kajian Pustaka	22
BAB III METODOLOGI PENELITIAN.....	26
3.1 Perancangan Algoritma	26

3.1.1	<i>Universal Sentence Encoder (USE)</i>	26
3.1.2	<i>Tensorflow Recommendation (TF-RS)</i>	28
3.2	Pengambilan Dataset	30
3.3	<i>Training Dataset</i>	31
3.4	<i>Loss Function</i>	32
3.5	<i>Similar Score</i>	33
3.6	<i>Deployment</i>	34
BAB IV	ANALISIS DAN HASIL.....	36
4.1	Tahap Persiapan Dataset	36
4.1.1	Analisis Penggunaan <i>Firebase</i>	37
4.2	Analisis Model <i>Universal Sentence Encoder (USE)</i>	40
4.2.1	<i>Analisis Similar Score</i>	46
4.3	Analisis Model <i>Tensorflow Recommendation (TFRS)</i>	48
4.3.1	<i>Analisis Loss Function</i>	54
4.4	Analisis Pengembangan Model dan Tingkat Akurasi	56
BAB IV	PENUTUP	59
5.1	Kesimpulan.....	59
5.2	Saran	60
DAFTAR	PUSTAKA	61
LAMPIRAN A	<i>LISTING CODE ALGORITHM</i>	66
LAMPIRAN B	HASIL REKOMENDASI	76
LAMPIRAN C	DATASET PENELITIAN.....	78

DAFTAR GAMBAR

Gambar 2. 1 Pola Kecerdasaan Buatan	7
Gambar 2. 2 <i>Machine Learning</i> Diagram	9
Gambar 2. 3 Daigram Alir NLP.....	15
Gambar 2. 4 Arsitektur <i>Word2Vec</i>	16
Gambar 2. 5 Arsitektur <i>Deep Averaging Network</i>	17
Gambar 2. 6 <i>Collaborative Filtering</i>	18
Gambar 2. 7 Vektor <i>Embedding</i>	20
Gambar 3.1 Rancangan Proses Algoritma USE.....	27
Gambar 3.2 Diagram Alir Proses Tensorflow Recommendation	29
Gambar 3.3 Tampilan <i>Website Mockaroo</i>	30
Gambar 3.4 <i>Training Process Model</i>	31
Gambar 3.5 Lokasi <i>File</i> Firebase.....	35
Gambar 4. 1 Dataset yang digunakan.....	36
Gambar 4. 2 Dataset <i>Job Interest</i>	37
Gambar 4. 3 Tampilan <i>Firebase</i>	38
Gambar 4. 4 Tampilan <i>JetAdmin</i>	39
Gambar 4. 5 Hasil Akhir Rekomendasi	46
Gambar 4. 6 Hasil Rekomendasi Metode TFRS	53
Gambar 4. 7 Lokasi File TFRS	54

DAFTAR TABEL

Tabel 4. 1 Hasil ID <i>User</i> dengan <i>Similar Score</i>	47
Tabel 4. 2 Hasil Pengembangan dan Tingkat Akurasi.....	57

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kondisi pengangguran di Indonesia mengalami peningkatan sejak pandemi COVID-19 melanda sejak awal tahun 2020. Menurut data Badan Pusat Statistik (BPS) pada Februari 2021, tingkat pengangguran terbuka di Indonesia sebesar 7,07% naik dari 5,50% pada Agustus 2020 [1]. Peningkatan pengangguran terutama disebabkan penurunan aktivitas ekonomi akibat pandemi. Banyaknya perusahaan yang mengalami kesulitan dan harus melakukan pemutusan hubungan kerja (PHK) atau pengurangan jam kerja. Gaya hidup masyarakat mulai berubah menjadi serba *online* dimana hal itu terjadi karena adanya perkembangan teknologi.

Meningkatnya penggunaan teknologi dalam kehidupan sehari-hari telah membawa dampak positif dan negatif, termasuk meningkatnya angka pengangguran, terutama di kalangan lulusan perguruan tinggi yang tinggi yang kurang memiliki *soft skill*, seperti kerja sama tim dan kolaborasi. Masalah ini adalah masalah serius yang ingin ditangani karena memiliki pengalaman yang sama. Untuk mengatasi masalah ini, sebagai hasil *capstone project* “Collabolio” hadir sebagai *platform* sosial-kolaboratif yang menggunakan algoritma berbasis perjodohan, yang biasanya digunakan pada aplikasi kencan, untuk kreasi portofolio diciptakan untuk membantu pengguna menemukan rekan kerja yang tepat untuk berkolaborasi dalam proyek-proyek kehidupan nyata yang bertujuan untuk mengatasi masalah pengembangan yang kurang berpengalaman terutama dalam kerja sama tim.

Platform ini memungkinkan individu yang memiliki ide proyek untuk menemukan pengembang dari berbagai jenis untuk membangun proyek mereka bersama-sama. Dengan demikian, Collabolio bertujuan untuk meningkatkan kemampuan kerja sama dan kolaborasi para lulusan perguruan tinggi dan menjawab pertanyaan tentang bagaimana *platform* sosial-kolaboratif yang menggunakan algoritma pencarian jodoh dapat membantu mengatasi masalah pengangguran di 10

kalangan lulusan perguruan tinggi di Indonesia, terutama terkait kurangnya kemampuan kerja sama dan kolaborasi.

Beberapa fitur utama Collabolio adalah *matchmaking*. *Matchmaking* adalah teknik yang digunakan untuk mencocokkan dua buah parameter atau lebih berdasarkan kesamaan makna atau tema yang terkandung dalam kalimat tersebut [2]. Pada *platform* ini menggunakan 2 buah parameter yaitu *skills* dan *job interest*. Model *machine learning* yang dipakai menggunakan metode *Universal Sentence Encoder* (USE) merupakan model *deep learning* yang dikembangkan oleh Google untuk numerik yang merepresentasikan makna dari kalimat tersebut. *Universal Sentence Encoder* mengkodekan teks ke dalam vektor dimensi tinggi yang dapat digunakan untuk klasifikasi teks, kemiripan semantik, pengelompokan, dan tugas-tugas bahasa alami lainnya [3].

Sebuah penelitian telah dilakukan untuk menggunakan *Universal Sentence Encoder* dalam mengklasifikasikan sentimen dari *tweet-tweet* yang berkaitan dengan COVID-19. Model ini berhasil mencapai tingkat akurasi sebesar 78,062%, yang lebih baik dibandingkan dengan klasifikasi mesin tradisional [4]. Penggunaan *Universal Sentence Encoder* lebih baik karena menggunakan pendekatan *transformer based deep learning* model. Selain itu USE memiliki keunggulan seperti mengkodekan makna dan semantik dari kalimat, memiliki generalisasi yang baik, dan kompatibilitas dengan berbagai bahasa [3].

Pada praktiknya, USE merupakan bagian dari *pipeline* pemrosesan bahasa, seperti dalam sistem rekomendasi, analisis teks, klasifikasi teks, dan tugas lainnya yang melibatkan pemahaman teks [5]. Pemrosesan bahasa alami memiliki tugas memahami dan menganalisis teks. Namun, untuk memahami makna keseluruhan teks perlu dilakukan representasi secara holistik. USE mampu mengatasi hal tersebut dengan menghasilkan representasi vektor secara keseluruhan. Representasi vektor ini disebut juga *embedding* atau *sentence embedding* [6].

Berdasarkan keunggulan tersebut USE menjadi algoritma terbaik yang cocok dalam permasalahan ini. Hal ini dikarenakan penggunaan beberapa algoritma lain seperti *Random Forest* untuk memprediksi program studi yang paling sesuai dengan profil calon mahasiswa baru merupakan upaya rintisan sistem rekomendasi

berbasis *machine learning*. Hasil penelitian yang menunjukkan bahwa model klasifikasi yang dikembangkan dapat memberikan rekomendasi program studi yang lebih akurat dibanding dengan sistem rekomendasi lainnya seperti *multinomial logistic regression* dan *support vector*. Dibuktikan dengan tingkat akurasi sebesar 86%, tingkat kepresisian model sebesar 84%, dan nilai *recall* sebesar 84%. Namun, *random forest* memiliki interpretabilitas yang terbatas karena menghasilkan model yang kompleks dan sulit untuk diinterpretasikan secara intuitif [7].

Stochastic gradient descent (SGD) memiliki performa terbaik dalam mengidentifikasi informasi *hoaks*, sementara *naïve bayes* memiliki performa paling rendah [8]. Sementara itu, dalam penelitian terkait rekomendasi, *cosine similarity* menghasilkan *precision* 100% dan *recall* 65% dengan *thresholding* 0.4, namun perlu perancangan metode yang lebih jelas untuk mempertimbangkan validitas dan reliabilitas akibat penggunaan ukuran sampel [5].

Metode *fuzzy inference system* (FIS) digunakan dalam penelitian untuk menghasilkan rekomendasi, dengan tingkat akurasi sebesar 63%. Meskipun FIS dapat menghasilkan tingkat akurasi yang lebih tinggi dari 61%, pengujian dalam penelitian ini dianggap belum maksimal karena hanya menggunakan 22 sampel data, sehingga hasilnya tidak dapat digeneralisasikan secara luas [9].

Tantangan dalam pengembangan sistem rekomendasi mencakup kekurangan data, keakuratan prediksi yang rendah, masalah privasi, integrasi model pembelajaran yang kompleks, dan faktor kontekstual yang lebih banyak [10]. Selain itu, masalah lainnya terjadi ketika sebagian besar *item* dalam dataset tidak memiliki *rating* atau *feedback* pengguna, menyebabkan kesulitan dalam memberikan rekomendasi yang akurat karena kurangnya informasi yang tersedia [11].

Greater Part Casting Voting Procedure dikenal sebagai *Majority Voting* atau *Plurality Voting* adalah teknik *ensemble learning* yang menggabungkan hasil dari beberapa model pembelajaran mesin untuk menghasilkan model yang akurat. Namun, kekurangan teknik ini adalah pengabaian bobot dalam pengambilan keputusan, yang dapat menghasilkan keputusan yang ambigu atau tidak pasti [12]. Pada proses rekomendasi, *cold start* terjadi saat sulit memberikan rekomendasi kepada pengguna baru tanpa data riwayat. Pengembangan teknik pengolahan

bahasa alami bertujuan untuk menangani pertumbuhan data dan menyediakan pengetahuan yang terstruktur [13]. *Tensorflow Recommenders* (TFRS) adalah *library* untuk membangun model sistem rekomendasi yang menangani masalah seperti *cold-start*, *multi-task learning*, dan pembelajaran yang tidak terawasi [14].

Transfer learning telah banyak digunakan dalam pengolahan bahasa alami melalui model bahasa yang telah dilatih secara mendalam, seperti *Bidirectional Encoder Representations from Transformers* dan *Universal Sentence Encoder* [15]. Model ini dilatih dan disesuaikan untuk pemrosesan teks multi-kata seperti kalimat, frasa, dan paragraf pendek. Pelatihannya melibatkan berbagai sumber data dan tugas untuk mengatasi berbagai tugas pemahaman bahasa alami dengan efektif [16].

Berdasarkan penelitian-penelitian sebelumnya terdapat beberapa kekurangan yaitu terdapat permasalahan *cold-start*, kekurangan data, keakuratan prediksi yang rendah, masalah privasi, integrasi model pembelajaran yang kompleks, dan interpretabilitas yang terbatas. Sehingga diperlukannya algoritma yang efektif untuk digunakan pada sistem rekomendasi yaitu *Universal Sentence Encoder* (USE) dan *Tensorflow Recommendation* (TFRS) menjadi algoritma terbaik untuk masalah saat ini karena pengimplementasiannya yang tidak begitu sulit untuk menangani kendala seperti *overfit* ataupun *underfit* dan dapat digunakan tanpa harus memahami parameter-parameter yang terlalu banyak.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada, maka terdapat permasalahan yang muncul dan dikerjakan pada penelitian, yaitu:

1. Bagaimana metode USE dan TFRS bekerja dalam melakukan sistem rekomendasi?
2. Bagaimana akurasi antara metode USE dengan metode TFRS?
3. Bagaimana metode USE dan TFRS meningkatkan akurasi?

1.3 Tujuan Penelitian

Pada penelitian yang dilakukan terdapat beberapa tujuan yang ingin dicapai, yaitu:

1. Mengetahui sistem rekomendasi menggunakan metode USE dan metode TFRS.
2. Mengetahui akurasi menggunakan metode USE dan metode TFRS.
3. Meningkatkan akurasi menggunakan metode USE dan metode TFRS.

1.4 Manfaat Penelitian

Sebuah penelitian dilakukan untuk mencapai suatu manfaat yang diharapkan, manfaat adanya penelitian ini, yaitu:

1. Memberikan wawasan tentang sistem rekomendasi menggunakan metode USE dan TFRS.
2. Memberikan pemahaman tentang kelebihan dan kekurangan masing-masing metode yaitu metode USE dan metode TFRS.
3. Pengembangan model menggunakan metode USE dan metode TFRS yang dapat menghasilkan peningkatan kinerja rekomendasi.

1.5 Batasan Masalah

Pada penelitian yang dilakukan terdapat beberapa batasan penelitian yang dapat dilihat sebagai berikut:

1. Algoritma dirancang menggunakan *keras*, *tensorflow*, dan *pandas*.
2. Dataset yang digunakan merupakan data yang di-generate pada *website mockaroo* dan terdapat dua buah parameter yang dijadikan acuan saat rekomendasi yaitu *job* dan *interest*.
3. Evaluasi performa dilakukan dengan menggunakan *Loss Function* untuk metode USE sedangkan untuk metode TFRS menggunakan *similar score*.
4. Komputasi menggunakan Google Colab dan Microsoft Excel.
5. Pada penelitian ini hanya sebatas mengembangkan model *machine learning*.

1.6 Sistematika Penulisan

Pada penulisan skripsi ini terdiri dari 5 bab. Isi setiap bab dapat diuraikan sebagai berikut.

BAB I Pendahuluan, pada bab ini menjelaskan mengenai latar belakang, tujuan penelitian, manfaat penelitian, tempat dan waktu penelitian, ruang lingkup penelitian, dan sistematika penulisan.

Selanjutnya,

BAB II Tinjauan Pustaka, pada bab ini menjelaskan tentang tinjauan pustaka yang berkaitan dengan penelitian ini.

BAB III Metodologi Penelitian, pada bab ini menjelaskan landasan teori yang digunakan.

BAB IV Hasil dan Analisis, pada bab ini berisi tentang hasil dari penelitian yang telah dilaksanakan beserta dengan analisis sesuai dengan batasan dan parameter yang digunakan.

BAB V Penutup, pada bab ini berisi kesimpulan dari penelitian yang dilakukan dan saran untuk pengembangan penelitian lebih lanjut.

DAFTAR PUSTAKA

- [1] B. P. Statistik, "Februari 2020: Tingkat Pengangguran Terbuka (TPT) sebesar 4,99 persen," Mei 2020.
- [2] I. K. Loukia Karanikola, "A Fuzzy Logic Approach for Reasoning under Uncertainty and Vagueness - A," in *2016 2nd International Conference on Information Management (ICIM)*, London, UK, 2016.
- [3] Y. Y. e. a. Cer Daniel, "Universal Sentence Encoder," *arXiv*, 2018.
- [4] I. E. A. K. F. E. W. Y. A. M. Fattoh, "Semantic Sentiment Classification for COVID-19 Tweets Using," *Hindawari*, 2022.
- [5] Y. C. D. A. A. G. M. L. J. C. N. K. R. Yang, "Multilingual Universal Sentence Encoder for Semantic Retrieval," *arXiv*, 2019.
- [6] N. & G. I. Reimers, "Sentence Embeddings using Siamese BERT-Networks," in *Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.
- [7] A. R. A. R. R. P. A. T. M. Pratama, "Model Klasifikasi Calon Mahasiswa Baru untuk Sistem Rekomendasi Program Studi Sarjana Berbasis Machine Learning," *Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 9, 2022.
- [8] S. M. N. Y. L. E. S. Tambunan, "Perbandingan Klasifikasi dengan Pendekatan Pembelajaran Mesin untuk Mengidentifikasi Tweet Hoaks di Media Sosial Twitter," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 7, 2021.
- [9] F. M. W. F. Y. C. Mar'i, "Sistem Rekomendasi Profesi Berdasarkan Dimensi Big Five Personality Menggunakan Fuzzy Inference System Tsukamoto," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 6, 2019.
- [10] S. M. S. F. H. Y. Nafea, "On Recommendation of Learning Object using Felder-Silverman Learning Style Model," *IEEE TRANSACTIONS and JOURNALS*, vol. 4, 2016.

- [11] A. A. R. A. T. L. S. M. X. Althbiti, "Addressing Data Sparsity in Collaborative Filtering Based Recommender System Using Clustering and Artificial Neural Network," *IEEE Computing and Communication Workshop and Conference*, 2021.
- [12] R. S. D. K. S. S., "Yoga Recommendation System for the Mental Well-Being of Students using Machine Learning," in *International Conference on Electronics, Communication and Aerospace Technology*, Coimbatore, India, 2022.
- [13] S. K. S. T. S. P. M. Tiwari, "Crop Recommendation using Machine Learning and Plant Disease Identification using CNN and Transfer-Learning Approach," in *IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India, 2022.
- [14] A. A, R. Y. P, S. R. Kumar, R. M and P. R, "Book Recommendation System with *Tensorflow*," in *International Conference on Advanced Computing & Communication System (ICACCS)*, Coimbatore, India, 2021.
- [15] M. N. S. J. I. P. Y. D. D. B. I. e. a. Bahri, "Implementasi Sistem Rekomendasi Makanan pada Aplikasi EatAja Menggunakan Algoritma Collaborative Filtering," *MULTINETICS*, vol. 7, 2021.
- [16] N. C. Y.-C. H. W.-L. Warikoo, "Lexically-aware Transformers based Bidirectional Encoder Representation Model for Learning Universal Bio-Entity Relations," Oxford University, 2020.
- [17] I. H. A. Amin, "Artificial Intelligence dalam Proses Industri Manufaktur," *Jurnal Teknologi Informasi DINAMIK*, vol. XIV, pp. 98-104, 2009.
- [18] T. Wahyudi, "Studi Kasus Pengembangan dan Penggunaan Artificial Intelligence (AI) Sebagai Penunjang Kegiatan Masyarakat Indonesia," *IJSE*, vol. 9, 2023.
- [19] U. Kalsum, "Pengenalan Kecerdasan Buatan (Artificial Intelligence) Kepada Para Remaja," *Jurnal Ilmiah Intech*, 2022.

- [20] J. W. G. Putra, "Pembelajaran Mesin dan Deep Learning," vol. 1.4.
- [21] R. K. Dinata and N. Hasdyna, *Machine Learning*, Sulawesi: Unimal Press, 2020.
- [22] S. Z. C. H. Z. J. Z. Sun, "A Survey of Optimization Methods from a Machine Learning Perspective," *IEEE Transactions on Cybernetics*, vol. 50, p. 8, 2020.
- [23] F. S. Singagerda, "Pengolahan Data Statistik dengan menggunakan program MiniTab," 2014.
- [24] N. Setiawan, *Pengolahan dan Analisis Data*, Bogor: Inspektorat Jendral Departemen Pendidikan Nasional, 2005.
- [25] J. W. G. Putra, *Pembelajaran Mesin dan Deep Learning*, Jan Wira Gotama Putra, 2020.
- [26] Python, "Tutorial Python," Python Software Foundation License Version 2, 06 July 2023. [Online]. Available: <https://docs.python.org/id/3.8/tutorial/>. [Accessed 17 January 2024].
- [27] GeeksforGeeks, GeeksforGeeks, 14 Apr 2023. [Online]. Available: <https://www.geeksforgeeks.org/introduction-deep-learning/>. [Accessed 18 Jan 2024].
- [28] J. N. F. M. W. C. A. W. A. Cristian, "Next Level Matchmaking," *Jurnal Mantik*, pp. 1-6, 2021.
- [29] A. Burchfiel, "Tokenex," 16 May 2022. [Online]. Available: <https://www.tokenex.com/blog/ab-what-is-nlp-natural-language-processing-tokenization/>. [Accessed 18 Jan 2024].
- [30] S. NLP, "Spark NLP," 29 July 2020. [Online]. Available: https://sparknlp.org/2020/07/29/lemma_id.html. [Accessed 18 Jan 2024].
- [31] GeeksforGeeks, "GeeksforGeeks," 03 Jan 2024. [Online]. Available: <https://www.geeksforgeeks.org/nlp-part-of-speech-default-tagging/>. [Accessed 18 Jan 2024].

- [32] N. Barney, "TechTarget," 2023. [Online]. Available: <https://www.techtarget.com/whatis/definition/named-entity-recognition-ner>. [Accessed 18 Jan 2024].
- [33] A. A. Awan, "DataCamp," Sept 2023. [Online]. Available: <https://www.datacamp.com/blog/what-is-tokenization>. [Accessed 18 Jan 2024].
- [34] A. J. Hanif, N. M. Farid and B. Hasanah, "Penerapan Natural Language Processing untuk Klasifikasi Bidang Minat berdasarkan Judul Tugas Akhir," *Jurnal Sistem Informasi dan Teknologi*, vol. 5, pp. 41-49, 2023.
- [35] A. Girsang, "Master of Computer Science," Binus, 17 November 2020. [Online]. Available: <https://mti.binus.ac.id/2020/11/17/word-embedding-dengan-word2vec/>. [Accessed 8 January 2024].
- [36] R. Maulid, "DQLab," 3 Oktober 2023. [Online]. Available: <https://dqlab.id/collaborative-filtering-pada-algoritma-data-science>. [Accessed 8 January 2024].
- [37] A. Chaudhary, "Amitness," 15 June 2020. [Online]. Available: <https://amitness.com/2020/06/universal-sentence-encoder/>. [Accessed 19 January 2024].
- [38] *Tensorflow*, "Tensorflow," 27 May 2023. [Online]. Available: https://www.tensorflow.org/recommenders/examples/basic_ranking. [Accessed 23 January 2024].
- [39] B. Shoumaker, "A Stata module for computing fertility rates and TFRs from birth histories: tfr2," *Demographic Research*, vol. 28, pp. 1093-1144, 2013.
- [40] Y. S. M. C. Z. Q. V. e. a. Wu, "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation," 2016. [Online]. Available: <https://research.google/pubs/googles-neural-machine-translation-system-bridging-the-gap-between-human-and-machine-translation/>. [Accessed 24 January 2024].

- [41] A. Najam, S. Latif and R. Latif, "Question Classification Using Universal Sentence Encoder and Deep Contextualized Transformer.," in *14th International Conference on Developments in eSystems Engineering (DeSE)*, 2021.
- [42] M. W. Putri, A. Muchayan and M. Kamisutara, "Sistem Rekomendasi Produk Pena Eksklusif Menggunakan Metode Contact-Base Filtering dan TF-IDF," vol. 2, pp. 229-236, JOINTECS.
- [43] K. S. Adi, K. Mustofa and I. A. Bustomi, "Pembuatan Model Matchmaking pada Permainan Hearthstone Menggunakan Hybrid dan Reciprocal Recommender System," 2019.
- [44] G. J. Lakshmi, S. A. N. Hilmi and A. Obaid, "Artificial intelligence agriculture recommendation model (AIARM)," *Internasional Journal of Health Sciences*, vol. 6, pp. 1782-1808, 2022.
- [45] Trivusi, "Mengenal 6 Jenis Loss Function pada Machine Learning," Trivusi, 17 September 2022. [Online]. Available: <https://www.trivusi.web.id/2022/08/loss-function.html>. [Accessed 20 Februari 2024].