

Implementasi Deep Factorization Machine Berbasis Profil Minat RIASEC untuk Rekomendasi Program Kepemudaan

Sihabuddin Rifqi*, Erna Dwi Astuti, Hidayatus Sibyan

Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Sains Al Qur'an, Indonesia

Email: ^{1,*}sihabuddinrifqi@mhs.unsiq.ac.id, ²ernada@unsiq.ac.id, ³hsibyan@unsiq.ac.id

Submitted: 05/06/2026; Accepted: 29/06/2026; Published: 30/06/2026

Abstrak—Pemilihan program kepemudaan yang kurang tepat akibat keterbatasan informasi dan pemetaan minat yang belum terstruktur menjadi kendala dalam pengembangan potensi pemuda. Penelitian ini bertujuan mengimplementasikan dan mengevaluasi efektivitas model *Deep Factorization Machine* (DeepFM) berbasis data profil minat instrumen RIASEC untuk menghasilkan rekomendasi program kepemudaan dari Kementerian Pemuda dan Olahraga (Kemenpora) yang terpersonalisasi. Penelitian kuantitatif eksperimental ini menggunakan *dataset* dari 485 responden yang dipetakan ke dalam lima program pemberdayaan. Fitur pengguna dibangun menggunakan vektor skor numerik enam dimensi RIASEC tanpa rekayasa fitur manual, yang diproses bersama fitur kategorikal program melalui integrasi komponen *Factorization Machine* dan *Deep Neural Network*. Hasil pengujian empiris menunjukkan arsitektur yang diusulkan memiliki performa prediktif yang sangat representatif. Sistem berhasil mencapai nilai *Normalized Discounted Cumulative Gain* (NDCG@3) sebesar 0,7064, *Area Under the Curve* (AUC) 0,7217, *Recall@3* 0,8018, dan *Precision@3* 0,4144. Capaian ini menunjukkan bahwa model DeepFM memberikan performa yang baik dalam menghasilkan serta menyusun peringkat 3 rekomendasi program teratas secara presisi sesuai dengan karakteristik psikometrik pengguna.

Kata Kunci: Sistem Rekomendasi; DeepFM; RIASEC; Program Kepemudaan; Deep Learning

Abstract—The inappropriate selection of youth programs due to limited information and unstructured interest mapping is a significant obstacle in developing youth potential. This study aims to implement and evaluate the effectiveness of the Deep Factorization Machine (DeepFM) model based on RIASEC interest profile data to generate personalized youth empowerment program recommendations from the Ministry of Youth and Sports (Kemenpora). This quantitative experimental research utilized a dataset of 488 respondents mapped into five empowerment programs. User features were constructed using the six-dimensional RIASEC numerical score vector without manual feature engineering, which was processed alongside categorical program features through the integration of Factorization Machine and Deep Neural Network components. Empirical testing results indicate that the proposed architecture possesses highly representative predictive performance. The system successfully achieved a Normalized Discounted Cumulative Gain (NDCG@3) score of 0.7064, Area Under the Curve (AUC) of 0.7217, Recall@3 of 0.8018, and Precision@3 of 0.4144. These achievements demonstrate that the DeepFM model delivers good performance in generating and ranking the top-3 program recommendations precisely according to the user's psychometric characteristics.

Keywords: Recommendation System; DeepFM; RIASEC; Youth Programs; Deep Learning

1. PENDAHULUAN

Pembangunan kualitas Sumber Daya Manusia (SDM) menjadi salah satu fokus utama dalam mendukung terwujudnya Visi Indonesia Emas 2045. Dalam konteks tersebut, pengembangan kapasitas pemuda menjadi agenda strategis nasional yang diwujudkan melalui berbagai program pemberdayaan yang diselenggarakan oleh Kementerian Pemuda dan Olahraga (Kemenpora) [1]. Program-program tersebut mencakup berbagai bidang, seperti kewirausahaan, kreativitas, pelatihan keterampilan, olahraga, hingga lingkungan hidup [2]. Namun, banyaknya pilihan program yang tersedia menyebabkan sebagian pemuda mengalami kesulitan dalam menentukan program yang sesuai dengan minat yang dimiliki. Keterbatasan informasi mengenai karakteristik program dan belum adanya mekanisme pemetaan minat yang terstruktur mengakibatkan pemuda berpotensi mengikuti program yang kurang sesuai dengan karakteristik dirinya [3]. Kondisi tersebut dapat berdampak pada kurang optimalnya manfaat program pemberdayaan yang diperoleh peserta serta menyulitkan proses pemetaan potensi pemuda secara lebih terarah [4]. Permasalahan tersebut menunjukkan pentingnya suatu pendekatan yang mampu mencocokkan profil minat individu dengan program yang tersedia secara lebih akurat dan berbasis data.

Salah satu pendekatan yang dipandang potensial untuk menjawab permasalahan tersebut adalah penerapan *deep learning* dengan model *Deep Factorization Machine* (DeepFM) pada data profil minat berbasis instrumen RIASEC. Instrumen RIASEC (*Realistic, Investigative, Artistic, Social, Enterprising, Conventional*) yang dikembangkan oleh Holland [5] mampu merepresentasikan kecenderungan minat seseorang ke dalam enam nilai numerik terukur, yang secara langsung dapat dimanfaatkan sebagai fitur masukan dalam proses pelatihan model. Model DeepFM yang diperkenalkan oleh Guo dkk. [6] bekerja dengan cara menggabungkan komponen *Factorization Machine* untuk menangkap interaksi antar fitur tingkat rendah dan komponen *Deep Neural Network* untuk mempelajari pola hubungan yang lebih kompleks secara bersamaan. Dengan mengumpukan skor dimensi RIASEC pengguna langsung ke dalam arsitektur DeepFM, model diharapkan dapat mempelajari pola kesesuaian antara karakteristik minat individu dan program kepemudaan secara lebih mendalam, tanpa bergantung pada data interaksi historis yang masif.

Beberapa penelitian terdahulu telah mengkaji penerapan model *deep learning* pada sistem rekomendasi maupun penggunaan instrumen RIASEC dalam pemetaan minat. Faqih dkk. [7] menguji pendekatan *Hybrid Filtering* dengan menggabungkan *Neural Collaborative Filtering* dan *Content-Based Filtering* pada data interaksi pengguna platform Steam, yang menghasilkan nilai RMSE sebesar 0,2211. Hasil tersebut menunjukkan bahwa model berbasis jaringan saraf mampu memodelkan preferensi pengguna dengan baik, meskipun pengujiannya masih bergantung pada data interaksi eksplisit yang melimpah di domain hiburan. Wen dan Geng [8] mengujicobakan model DeepFM yang disempurnakan pada data rekomendasi pendidikan anak usia dini dan memperoleh nilai AUC sebesar 0,9341, membuktikan bahwa arsitektur DeepFM mampu menangkap pola minat pengguna secara akurat bahkan pada data yang berdimensi tinggi dan bersifat jarang. Hidayat dkk. [9] menguji model *Hybrid Neural Network* dan *Collaborative Filtering* berbasis profil RIASEC untuk memprediksi kesesuaian program studi bagi siswa SMK, dengan hasil *precision* 87% dan *recall* 82%, yang mengonfirmasi bahwa skor dimensi RIASEC layak digunakan sebagai representasi fitur pengguna dalam model prediktif. Sulistyio dkk. [10] menganalisis validitas Teori Holland dalam konteks bimbingan karier dan menegaskan bahwa keenam dimensi RIASEC memiliki dasar psikometrik yang kuat untuk dioperasionalkan sebagai nilai numerik terukur. Zhang [11] menjelaskan secara mendalam landasan teoritis *Matrix Factorization* dan *Factorization Machine* serta keterkaitannya dengan algoritma *deep learning* dalam sistem rekomendasi, termasuk cara kerja komponen FM dalam arsitektur DeepFM melalui pendekatan perhitungan bertahap yang mudah dipahami. Kajian tersebut memperkuat pemahaman bahwa FM mampu memodelkan interaksi antar fitur secara efisien, yang menjadi dasar mengapa komponen ini relevan untuk diintegrasikan dalam model DeepFM pada penelitian ini.

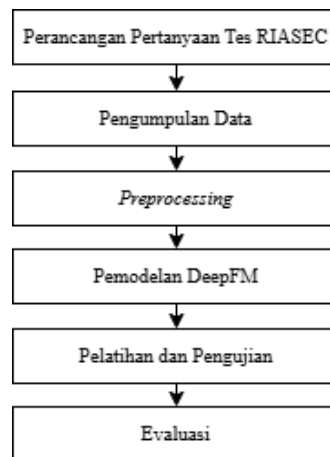
Mencermati penelitian-penelitian tersebut, tampak bahwa pengujian model *deep learning* untuk rekomendasi telah menunjukkan perkembangan yang cukup signifikan. Namun demikian, jika ditelaah lebih jauh, sebagian besar pengujian yang dilakukan masih bertumpu pada data interaksi historis pengguna seperti riwayat klik atau pembelian sebagai sumber utama fitur [7], sehingga kemampuan model dalam memproses data psikometrik yang bersifat *dense* dan kontinu seperti skor RIASEC belum banyak dieksplorasi. Di sisi lain, penelitian yang sudah menggunakan RIASEC [8], [9], [10] sebagai representasi pengguna umumnya menguji model yang lebih sederhana, sehingga sejauh mana arsitektur yang lebih dalam seperti DeepFM mampu mengekstraksi pola interaksi yang lebih kompleks dari dimensi-dimensi minat individu masih menjadi pertanyaan yang belum terjawab. Penerapan DeepFM secara khusus pada data profil RIASEC dalam konteks program kepemudaan pun tampaknya belum pernah dikaji sebelumnya. Berdasarkan kesenjangan tersebut, penelitian ini hadir untuk mengujicobakan arsitektur DeepFM menggunakan vektor skor RIASEC sebagai fitur masukan utama, untuk melihat seberapa efektif model tersebut dalam mempelajari pola kesesuaian antara karakteristik minat pemuda dan program yang paling sesuai bagi mereka.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk menguji dan menganalisis efektivitas model DeepFM dalam memproses data profil minat RIASEC untuk menghasilkan prediksi kesesuaian program kepemudaan Kemendikbud yang terpersonalisasi. Pengujian dilakukan secara empiris menggunakan data dari 488 responden yang dipetakan terhadap lima program pemberdayaan, dengan evaluasi menggunakan metrik AUC, *Precision@K*, *Recall@K*, dan *NDCG@K* [12] guna mengukur kemampuan diskriminasi, ketepatan, kelengkapan, serta kualitas *ranking* rekomendasi yang dihasilkan model secara menyeluruh. Penelitian ini diharapkan dapat memberikan bukti empiris mengenai kelayakan penggunaan data psikometrik RIASEC sebagai fitur masukan dalam arsitektur DeepFM, sekaligus membuka peluang eksplorasi lebih lanjut mengenai penerapan model *deep learning* pada data berbasis profil minat di berbagai konteks pengembangan SDM.

2. METODOLOGI PENELITIAN

2.1 Jenis dan Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksperimental berbasis *deep learning*, yaitu pendekatan yang berfokus pada pengujian dan pengukuran performa suatu model secara objektif melalui data numerik dan metrik evaluasi yang telah ditentukan [13]. Pendekatan eksperimental ini dipilih karena penelitian melibatkan proses perancangan arsitektur model, pelatihan menggunakan data numerik, pengujian performa, serta analisis hasil prediksi secara objektif dan terukur. Dalam pendekatan ini, peneliti merancang dan membangun arsitektur DeepFM yang menggabungkan komponen *Factorization Machine* dan *Deep Neural Network*, kemudian mengamati pengaruh penerapan arsitektur tersebut terhadap kualitas prediksi kesesuaian program berdasarkan vektor skor RIASEC pengguna sebagai fitur masukan utama. Melalui pendekatan kuantitatif eksperimental ini, dimungkinkan dilakukannya pengujian kinerja model secara sistematis berdasarkan konfigurasi parameter yang telah ditentukan, seperti jumlah lapisan tersembunyi, ukuran *embedding*, dan laju pembelajaran. Selain itu, pendekatan ini memungkinkan analisis mendalam terhadap kemampuan model dalam mengenali pola kesesuaian antara karakteristik minat individu yang bersifat psikometrik dengan program kepemudaan yang tersedia. Dengan demikian, penggunaan pendekatan kuantitatif eksperimental dinilai tepat untuk menjawab tujuan penelitian, yaitu menguji efektivitas model DeepFM dalam memproses data profil minat RIASEC serta mengevaluasi kinerjanya berdasarkan metrik AUC, *Precision@K*, *Recall@K*, dan *NDCG@K* yang jelas dan terukur. Adapun tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.2 Sumber Data

Penelitian ini menggunakan dua jenis sumber data. Data primer diperoleh melalui instrumen tes minat berbasis metode RIASEC yang disebarakan kepada responden. Responden merupakan pemuda yang pernah mengikuti program pemberdayaan Kemenpora atau program serupa, sehingga memiliki pengalaman nyata yang dapat merepresentasikan keterkaitan antara profil minat individu dan program yang pernah diikuti. Data hasil tes berupa vektor skor numerik pada keenam dimensi RIASEC yang selanjutnya digunakan sebagai fitur masukan (*input features*) dalam pelatihan model DeepFM.

Data sekunder diperoleh dari informasi program pemberdayaan yang diselenggarakan oleh Kemenpora, yang dalam penelitian ini dibatasi pada lima program sebagai label target, yaitu Wirausaha Muda Pemula (WMP), Kreativesia, Pelatihan Tenaga Penggerak Olahraga Nasional (TPON), *Youth Life Skills*, dan Pemuda Peduli Lingkungan Asri dan Bersih (Pepelingsasih) [14]. Kelima program tersebut dipilih karena masing-masing memiliki karakteristik yang berbeda dan dapat dipetakan terhadap dimensi-dimensi RIASEC tertentu, sehingga membentuk ruang label yang cukup beragam untuk melatih dan menguji kemampuan diskriminasi model. Kombinasi antara data primer berupa profil RIASEC pengguna dan data sekunder berupa karakteristik program inilah yang membentuk *dataset* utama yang digunakan dalam seluruh tahapan eksperimen penelitian ini.

2.3 Metode RIASEC

Penelitian ini menggunakan metode RIASEC yang dikembangkan oleh John L. Holland sebagai dasar pemetaan profil minat pengguna. Teori Holland menyatakan bahwa kecenderungan minat individu dapat dikelompokkan ke dalam enam tipe kepribadian, yaitu *Realistic* (R), *Investigative* (I), *Artistic* (A), *Social* (S), *Enterprising* (E), dan *Conventional* (C). Masing-masing dimensi merepresentasikan kecenderungan aktivitas, ketertarikan, dan karakteristik individu yang berbeda [5], sebagaimana dirangkum pada Tabel 1.

Tabel 1. Deskripsi Dimensi RIASEC

Dimensi	Kode	Karakteristik
<i>Realistic</i>	R	Menyukai aktivitas fisik, teknis, dan penggunaan alat atau mesin
<i>Investigative</i>	I	Menyukai kegiatan analitis, ilmiah, dan pemecahan masalah kompleks
<i>Artistic</i>	A	Menyukai ekspresi kreatif, seni, dan lingkungan yang tidak terstruktur
<i>Social</i>	S	Menyukai interaksi sosial, membantu, dan mendidik orang lain
<i>Enterprising</i>	E	Menyukai kepemimpinan, persuasi, dan kegiatan berorientasi bisnis
<i>Conventional</i>	C	Menyukai pekerjaan terstruktur, administratif, dan berorientasi data

Dalam penelitian ini, instrumen RIASEC diimplementasikan dalam bentuk kuesioner yang terdiri dari sejumlah pernyataan yang telah disesuaikan dengan konteks program pemberdayaan pemuda. Setiap pernyataan dirancang untuk mengukur kecenderungan responden pada salah satu dari keenam dimensi tersebut. Hasil pengisian kuesioner menghasilkan enam nilai skor numerik yang merepresentasikan tingkat kecenderungan responden pada masing-masing dimensi. Keenam nilai skor inilah yang kemudian digunakan sebagai vektor fitur masukan (*input feature vector*) dalam proses pelatihan model DeepFM, menggantikan pendekatan konvensional yang umumnya menggunakan data interaksi historis pengguna [15]. Representasi profil pengguna dinyatakan sebagai (1):

$$u = [R, I, A, S, E, C] \tag{1}$$

dengan R, I, A, S, E, C menyatakan skor masing-masing dimensi RIASEC.

2.4 Representasi Pengguna dan Program

Dalam penelitian ini, setiap pengguna direpresentasikan sebagai vektor fitur numerik yang diperoleh langsung dari hasil tes RIASEC pada tahap sebelumnya. Vektor tersebut terdiri dari enam nilai skor yang mencerminkan tingkat kecenderungan individu pada masing-masing dimensi minat, sebagaimana telah didefinisikan pada persamaan (2). Representasi ini berfungsi sebagai *user features* yang diumpangkan ke dalam arsitektur DeepFM untuk proses pembelajaran pola kesesuaian.

Di sisi item, setiap program pemberdayaan Kemenpora direpresentasikan sebagai vektor *item features* yang memuat informasi karakteristik masing-masing program. Kelima program yang digunakan dalam penelitian ini masing-masing memiliki profil karakteristik yang berbeda dan dapat dipetakan ke dimensi RIASEC tertentu. Representasi program ke- j dinyatakan sebagai (2):

$$p_j = [p_{j1}, p_{j2}, \dots, p_{jm}] \tag{2}$$

dengan p_j merupakan vektor fitur program pemberdayaan ke- j dan m adalah jumlah fitur yang merepresentasikan karakteristik program tersebut.

Proses pembelajaran model DeepFM dilakukan dengan memprediksi skor preferensi numerik $\hat{y}(u, p_j)$ untuk setiap pasangan pengguna-program, yang merepresentasikan seberapa besar kesesuaian antara profil minat pengguna dengan program p_j . Skor preferensi ini diformulasikan sebagai (3):

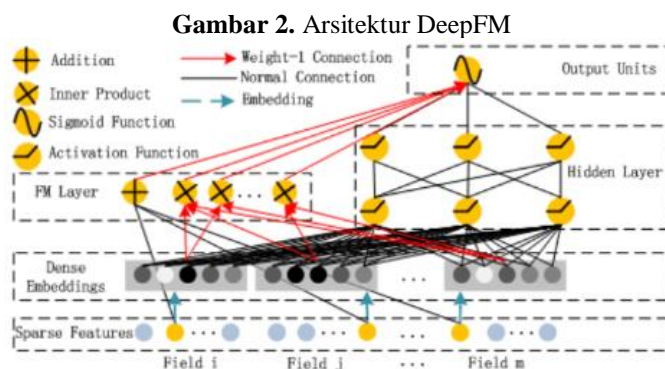
$$\hat{y}(u, p_j) = \text{DeepFM}(u, p_j) \tag{3}$$

Skor preferensi yang dihasilkan kemudian digunakan untuk menyusun peringkat (*ranking*) seluruh program yang tersedia bagi setiap pengguna, dari program yang paling relevan hingga yang paling kurang sesuai. Dengan demikian, output akhir model bukan berupa satu label kelas tunggal, melainkan berupa urutan program yang terpersonalisasi berdasarkan besarnya skor preferensi yang diprediksi untuk masing-masing pasangan pengguna-program.

2.5 Model DeepFM

Penelitian ini menggunakan model *Deep Factorization Machine* (DeepFM) sebagai metode utama dalam memproses data profil minat RIASEC untuk menghasilkan skor preferensi antara pengguna dan program kepemudaan. DeepFM merupakan arsitektur yang menggabungkan komponen *Factorization Machine* (FM) dan *Deep Neural Network* (DNN) dalam satu kerangka pembelajaran yang terintegrasi [6]. Keunggulan utama DeepFM terletak pada kemampuannya mempelajari interaksi fitur tingkat rendah (*low-order feature interaction*) melalui komponen FM dan hubungan non-linear yang lebih kompleks (*high-order feature interaction*) melalui komponen DNN secara bersamaan dalam satu proses pelatihan [8]. Dibandingkan dengan metode yang lebih konvensional, DeepFM tidak memerlukan rekayasa fitur (*feature engineering*) secara manual karena proses pembelajaran interaksi antar fitur dilakukan secara otomatis oleh model [16].

Pada penelitian ini, vektor skor RIASEC pengguna dan vektor fitur program pemberdayaan diumpangkan secara langsung sebagai masukan model DeepFM. Model kemudian mempelajari pola interaksi laten antara kedua representasi tersebut untuk menghasilkan skor preferensi numerik yang mencerminkan tingkat kesesuaian antara profil minat pengguna dengan masing-masing program. Arsitektur DeepFM yang diterapkan dalam penelitian ini digambarkan pada Gambar 2.



Arsitektur DeepFM terdiri dari dua komponen utama yang bekerja secara paralel, yaitu komponen FM dan komponen DNN, yang masing-masing menerima masukan yang sama namun memproses informasi dengan cara yang berbeda.

Komponen FM digunakan untuk mempelajari interaksi antar fitur tingkat rendah (*low-order*), termasuk interaksi orde pertama dan orde kedua antar dimensi fitur. Komponen FM dinyatakan sebagai (4):

$$y_{FM} = \langle w, x \rangle + \sum_{i=1}^n \sum_{j=i+1}^n (v_i, v_j) x_i x_j \quad (4)$$

dengan w merupakan vektor bobot interaksi orde pertama, x adalah vektor fitur masukan, v_i dan v_j merupakan vektor laten dari fitur ke- i dan ke- j , serta n adalah jumlah total fitur masukan.

Komponen DNN digunakan untuk mempelajari hubungan non-linear yang lebih kompleks (*high-order feature interaction*) antar fitur melalui beberapa lapisan tersembunyi (*hidden layers*). Keluaran komponen DNN dinyatakan sebagai (5):

$$y_{DNN} = \sigma(W^{(l)} \cdot a^{(l-1)} + b^{(l)}) \quad (5)$$

dengan $W^{(l)}$ dan $b^{(l)}$ masing-masing merupakan matriks bobot dan bias pada lapisan ke- l , $a^{(l-1)}$ adalah keluaran lapisan sebelumnya, dan σ adalah fungsi aktivasi.

Skor preferensi akhir diperoleh dari kombinasi keluaran kedua komponen melalui fungsi aktivasi sigmoid, yang memastikan nilai keluaran berada pada rentang 0-1 (6):

$$\hat{y} = \sigma(y_{FM} + y_{DNN}) \quad (6)$$

dengan \hat{y} merupakan skor preferensi yang diprediksi model untuk pasangan pengguna-program tertentu. Nilai \hat{y} yang lebih tinggi mengindikasikan tingkat kesesuaian yang lebih besar antara profil minat pengguna dengan program yang bersangkutan, dan selanjutnya digunakan sebagai dasar penyusunan peringkat rekomendasi.

2.6 Evaluasi Model

Evaluasi dilakukan untuk mengukur kemampuan model DeepFM dalam menghasilkan peringkat program pemberdayaan yang sesuai dengan profil minat pengguna. *Dataset* dibagi menjadi data pelatihan (*training data*), data validasi (*validation data*), dan data pengujian (*testing data*) dengan proporsi 70:15:15. Proporsi ini dipilih untuk memberikan alokasi data yang cukup besar (70%) agar model *deep learning* dapat mempelajari pola interaksi secara optimal, sekaligus menyediakan porsi yang seimbang (masing-masing 15%) untuk keperluan validasi parameter dan evaluasi objektif [17]. Selain itu, proses pembagian ini menerapkan pendekatan *user-based split*, di mana keseluruhan data interaksi dari satu pengguna unik dikelompokkan secara utuh ke dalam satu himpunan data. Pendekatan ini diterapkan secara spesifik untuk mencegah terjadinya kebocoran data (*data leakage*) dan memastikan bahwa model dievaluasi secara murni berdasarkan kemampuannya merekomendasikan program kepada pengguna baru (*unseen users*) [18].

Model dievaluasi menggunakan empat metrik utama, yaitu AUC, Precision@K, Recall@K, dan NDCG@K, yang dipilih karena sesuai dengan karakteristik *output* model berupa skor preferensi numerik untuk keperluan *ranking* [17]. Nilai K merepresentasikan batasan jumlah daftar rekomendasi teratas yang dievaluasi dan disajikan oleh sistem kepada pengguna, yang pada penelitian ini ditetapkan sebesar 3.

Area Under the Curve (AUC) digunakan untuk mengukur kemampuan model dalam membedakan antara program yang relevan dan tidak relevan bagi pengguna. AUC merepresentasikan probabilitas bahwa model akan memberikan skor prediksi lebih tinggi pada pasangan pengguna-program yang relevan dibandingkan yang tidak relevan. Nilai AUC dihitung menggunakan (7):

$$AUC = \frac{\sum_{i \in \text{positif}} \text{rank}_i - \frac{n_{\text{pos}}(n_{\text{pos}}+1)}{2}}{n_{\text{pos}} \times n_{\text{neg}}} \quad (7)$$

dengan n_{pos} dan n_{neg} masing-masing adalah jumlah sampel positif dan negatif, serta rank_i adalah peringkat sampel positif ke- i dalam daftar yang diurutkan berdasarkan skor prediksi [19].

Precision@K digunakan untuk mengukur proporsi program yang direkomendasikan dan benar-benar relevan terhadap seluruh program yang masuk dalam K daftar rekomendasi teratas [20]. Nilai *Precision@K* dihitung sebagai (8):

$$\text{Precision}@k = \frac{\text{jumlah item relevan dalam } k \text{ rekomendasi teratas}}{k} \quad (8)$$

Recall@K digunakan untuk mengukur proporsi program relevan yang berhasil ditemukan oleh model pada K urutan teratas dibandingkan dengan seluruh program relevan yang faktualnya pernah diikuti atau sesuai bagi pengguna [21]. Nilai *Recall@K* dihitung menggunakan (9):

$$Recall@k = \frac{\text{jumlah item relevan dalam } k \text{ rekomendasi teratas}}{\text{total item relevan}} \quad (9)$$

Normalized Discounted Cumulative Gain (NDCG@K) digunakan untuk mengevaluasi kualitas urutan (*ranking*) dari *K* daftar rekomendasi teratas yang dihasilkan model. Metrik ini memberikan bobot lebih tinggi pada program relevan yang muncul di posisi paling atas dalam daftar rekomendasi. Nilai NDCG@K dihitung menggunakan (10):

$$NDCG@K = \frac{DCG@K}{IDCG@K} \quad (10)$$

dengan nilai *Discounted Cumulative Gain* (DCG@K) dihitung sebagai (11):

$$DCG@K = \sum_{i=1}^K \frac{rel_i}{\log_2(i+1)} \quad (11)$$

di mana rel_i merupakan skor relevansi item pada posisi ke- i dan IDCG@K merupakan nilai DCG@K ideal yang diperoleh dari urutan rekomendasi yang sempurna [22]. Semakin tinggi nilai AUC, *Precision@K*, *Recall@K*, dan NDCG@K, maka performa model dalam menghasilkan peringkat rekomendasi yang akurat dan relevan dinilai semakin baik [23], [24].

3. HASIL DAN PEMBAHASAN

3.1 Perancangan Instrumen dan Dataset

Instrumen pengukuran minat disusun menggunakan pendekatan metode RIASEC yang dikembangkan berdasarkan teori Holland. Pertanyaan pada instrumen diadaptasi dari ONET Interest Profiler [25] dengan sejumlah penyesuaian terhadap konteks program pemberdayaan pemuda di Indonesia, agar setiap pertanyaan lebih relevan dengan aktivitas pengembangan diri, kewirausahaan, kreativitas, olahraga, dan kegiatan sosial yang berkaitan dengan program Kemenpora. Instrumen final terdiri dari 42 pertanyaan yang terbagi secara merata ke dalam enam dimensi RIASEC, dengan masing-masing dimensi diwakili oleh tujuh pertanyaan.

Responden memberikan jawaban menggunakan skala *Likert* dengan rentang nilai 1-5, di mana nilai 1 menunjukkan "sangat tidak sesuai" dan nilai 5 menunjukkan "sangat sesuai" terhadap pernyataan yang diberikan. Skor pada setiap dimensi RIASEC diperoleh dari akumulasi jawaban tujuh pertanyaan pada dimensi yang bersangkutan, sehingga setiap dimensi memiliki rentang skor 7-35. Keenam nilai skor inilah yang membentuk vektor fitur pengguna sebagaimana didefinisikan pada persamaan (1), dan selanjutnya digunakan sebagai fitur masukan dalam proses pelatihan model DeepFM.

Sebelum instrumen digunakan untuk pengumpulan data secara luas, dilakukan uji validitas dan reliabilitas untuk memastikan kelayakan alat ukur. Uji validitas dilakukan menggunakan metode korelasi *Pearson Product Moment* terhadap 42 butir pernyataan. Hasil pengujian menunjukkan bahwa seluruh butir pernyataan memiliki nilai signifikansi ($p\text{-value}$) < 0,05 dan nilai korelasi item-total lebih besar dibandingkan nilai $r\text{-tabel}$ pada taraf signifikansi 5%, sehingga seluruh item dinyatakan valid dan mampu mengukur konstruk minat yang digunakan. Selanjutnya, uji reliabilitas dievaluasi menggunakan koefisien *Cronbach's Alpha* untuk mengetahui tingkat konsistensi internal instrumen. Hasil perhitungan menunjukkan nilai *Cronbach's Alpha* sebesar 0,935. Nilai tersebut lebih besar dari batas minimum reliabilitas 0,60, sehingga instrumen tes minat RIASEC dinyatakan sangat reliabel [26]. Hal ini menunjukkan bahwa seluruh pernyataan memiliki konsistensi yang tinggi dalam mengukur karakteristik minat responden dan layak digunakan sebagai fitur masukan pada model sistem rekomendasi [27].

Pengumpulan data dilakukan secara daring menggunakan kuesioner digital yang disebarakan hingga mendapatkan 485 responden. Responden merupakan pemuda yang pernah mengikuti program pemberdayaan Kemenpora maupun program serupa, sehingga setiap responden memiliki rekam jejak keterlibatan program yang dapat dijadikan dasar pembentukan label data. Setiap responden kemudian dipetakan ke dalam salah satu atau lebih dari lima kategori program pemberdayaan yang menjadi label target dalam penelitian ini, yaitu WMP, Kreativesia, TPON, *Youth Life Skills*, dan Pepelingsih.

Dataset yang terbentuk terdiri dari tiga komponen utama, yaitu data profil pengguna, vektor skor enam dimensi RIASEC, dan label program pemberdayaan yang pernah diikuti oleh masing-masing responden. Struktur *dataset* ini dirancang agar model DeepFM dapat mempelajari pola kesesuaian antara profil minat pengguna dan karakteristik program secara lebih menyeluruh.

3.2 Preprocessing dan Representasi Data

Tahap *preprocessing* dilakukan untuk mempersiapkan *dataset* sebelum digunakan pada proses pelatihan model DeepFM. Pada tahap ini dilakukan pemeriksaan data kosong (*missing value*), normalisasi data numerik, transformasi kategori program, serta pembentukan relasi antara pengguna dan program. Proses ini bertujuan agar data memiliki format yang sesuai dengan kebutuhan model rekomendasi berbasis *deep learning*.

Skor hasil tes RIASEC yang diperoleh dari masing-masing responden digunakan sebagai fitur numerik (*dense features*). Setiap pengguna direpresentasikan menggunakan enam nilai utama, yaitu *Realistic* (R), *Investigative* (I), *Artistic* (A), *Social* (S), *Enterprising* (E), dan *Conventional* (C). Selain itu, data program pemberdayaan dikonversi menjadi fitur kategori (*sparse features*) menggunakan proses *label encoding* sehingga setiap program memiliki representasi numerik yang dapat diproses oleh model.

Pada tahap pembentukan *dataset* interaksi (*user-item interaction*), setiap baris data merepresentasikan hubungan antara satu pengguna dan satu program pemberdayaan. Nilai label digunakan untuk menunjukkan apakah pengguna memiliki keterkaitan terhadap program tertentu berdasarkan data faktual responden (direpresentasikan sebagai label 1). Untuk melatih model klasifikasi biner, *dataset* diperluas menggunakan teknik *negative sampling*. Artinya, untuk setiap program yang pernah diikuti responden (label 1), sistem juga memasukkan sampel program yang tidak diikuti oleh responden tersebut dan melabelinya sebagai data negatif (label 0) sebagaimana pada Tabel 2. Dengan pendekatan ini, model dapat membedakan pola hubungan secara lebih tajam antara karakteristik minat pengguna dengan program yang relevan maupun yang tidak relevan. Pada akhir tahap ini, proses perluasan tersebut menghasilkan *dataset* interaksi utuh berjumlah 2.425 baris, yang terdiri dari 759 sampel positif (label 1) dan 1.666 sampel negatif (label 0).

Tabel 2. Tabel *User-Item Interaction*

ID	Usia	JK	R	I	A	S	E	C	Program	item	label
1	22	L	29	29	26	30	28	19	Inovasi/Kesenian, Soft Skill	WMP	0
1	22	L	29	29	26	30	28	19	Inovasi/Kesenian, Soft Skill	Youth Life Skills	1
1	22	L	29	29	26	30	28	19	Inovasi/Kesenian, Soft Skill	Kreativesia	1
1	22	L	29	29	26	30	28	19	Inovasi/Kesenian, Soft Skill	TPON	0
1	22	L	29	29	26	30	28	19	Inovasi/Kesenian, Soft Skill	Pepelingsih	0

Dataset kemudian dibagi menggunakan pendekatan *user-based split* menjadi tiga bagian, yaitu data pelatihan (*training data*), data validasi (*validation data*), dan data pengujian (*testing data*) dengan proporsi 70:15:15 dengan rincian jumlah data seperti pada Tabel 3. Data pelatihan digunakan untuk proses pembelajaran model DeepFM, data validasi digunakan untuk memantau performa model selama proses pelatihan dan membantu mengurangi risiko *overfitting*, sedangkan data pengujian digunakan untuk mengevaluasi performa akhir sistem dalam menghasilkan rekomendasi program pemberdayaan yang relevan terhadap profil pengguna.

Tabel 3. Hasil Pembagian Data

Data	Jumlah Baris	Label 1	Label 0	Jumlah User
<i>Training</i>	1.695	532	1.163	339
<i>Validation</i>	360	109	251	72
<i>Testing</i>	370	118	252	74
Total	2.425	759	1.666	485

3.3 Perancangan Model DeepFM

Perancangan model dilakukan dengan menggunakan arsitektur DeepFM yang menggabungkan komponen FM dan DNN. Pada penelitian ini, fitur pengguna yang digunakan meliputi usia dan vektor skor pada keenam dimensi RIASEC. Selain itu, data program pemberdayaan diikutsertakan sebagai fitur kategorikal (*sparse features*) dalam sistem rekomendasi. Fitur kategorikal tersebut diproses menggunakan teknik *embedding* sehingga setiap kategori direpresentasikan ke dalam bentuk vektor padat berdimensi rendah (*low-dimensional dense vector*). Representasi *embedding* ini memungkinkan model mempelajari hubungan laten antar karakteristik program secara lebih efisien dibandingkan representasi kategorikal tradisional.

Dalam arsitektur yang dibangun, komponen FM difokuskan untuk mempelajari interaksi fitur tingkat rendah (*low-order feature interactions*). Sementara itu, komponen DNN digunakan untuk memodelkan interaksi fitur non-linear yang memiliki kompleksitas tinggi (*high-order feature interactions*). Untuk menangkap pola preferensi pemuda yang kompleks dan mendalam, arsitektur DNN pada penelitian ini dirancang menggunakan tiga lapisan tersembunyi (*hidden layer*) dengan jumlah neuron yang menurun secara bertingkat, yaitu 256, 128, dan 64 neuron. Struktur hierarkis ini memungkinkan model untuk mengekstraksi representasi fitur secara bertahap, mulai dari tingkat yang paling umum hingga pola yang sangat spesifik terkait kesesuaian minat pengguna.

Guna menjaga stabilitas pembelajaran dan mencegah terjadinya *overfitting* akibat penggunaan arsitektur yang dalam (*deep*), model ini dioptimasi dengan mekanisme regularisasi yang komprehensif. Regularisasi tersebut meliputi penerapan *dropout* sebesar 0,5 yang berfungsi mereduksi dependensi antar neuron secara acak pada saat pelatihan, serta penalti regularisasi L2 (*weight decay*) sebesar 0,01 yang diaplikasikan baik pada *hidden layer* DNN maupun *embedding layer*. Kombinasi parameter ini terbukti krusial dalam menjaga keseimbangan model agar tetap memiliki kapasitas pembelajaran yang tinggi tanpa kehilangan kemampuan generalisasinya terhadap data

baru. Pada tahap akhir, keluaran dari komponen FM dan DNN digabungkan untuk menghasilkan skor prediksi numerik yang merepresentasikan tingkat kecocokan pengguna terhadap program pemberdayaan yang tersedia..

3.4 Optimasi dan Pelatihan Model

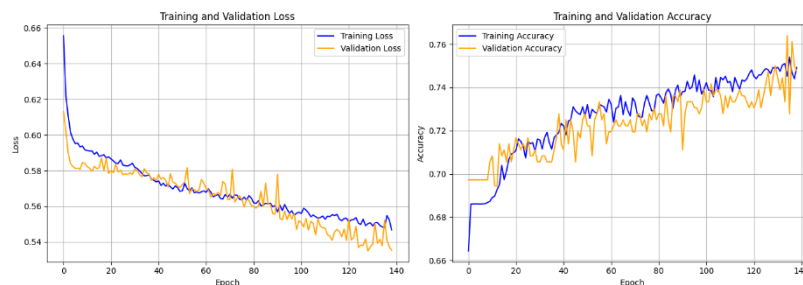
Proses optimisasi dan pelatihan bertujuan agar model DeepFM dapat mempelajari bobot terbaik dalam memprediksi tingkat kecocokan antara karakteristik pemuda dan program pemberdayaan. Pada penelitian ini, model dikompilasi menggunakan algoritma optimasi Adam dengan *learning rate* sebesar 0,0005. Algoritma ini dipilih karena kemampuannya dalam melakukan pembaruan bobot secara adaptif. Fungsi kerugian (*loss function*) yang digunakan adalah *binary crossentropy*, yang dirancang khusus untuk meminimalkan tingkat kesalahan pada kasus prediksi probabilitas biner.

Sebelum mencapai konfigurasi pelatihan yang optimal, penelitian ini melakukan serangkaian evaluasi penyesuaian parameter (*hyperparameter tuning*) untuk mengatasi masalah stagnasi pembelajaran (*underfitting*) dan penghafalan data (*overfitting*). Berdasarkan hasil uji coba empiris pada data berskala moderat ini, keseimbangan model terbaik (*good fit*) tercapai dengan mempertahankan arsitektur DNN yang dalam (256, 128, dan 64 neuron) namun diimbangi dengan regularisasi yang ketat. Penggunaan *dropout* sebesar 0,5 dan regularisasi L2 sebesar 0,01 terbukti krusial dalam memaksa model mengenali pola esensial tanpa harus kehilangan kapasitas generalisasinya terhadap data baru.

Pelatihan model kemudian dikonfigurasi untuk berjalan hingga maksimal 150 *epoch* dengan ukuran *batch* (*batch size*) sebesar 32. Penggunaan *batch size* 32 dipilih untuk memberikan efek regulasi alami berupa sedikit fluktuasi (*noise*) pada saat pembaruan gradien (*gradient descent*), yang secara efektif membantu model untuk tidak terjebak pada area minimum lokal (*local minima*). Untuk memantau performa model pada data baru secara akurat, penelitian ini mengimplementasikan fungsi *callback* kustom yang bertugas menghitung metrik *validation loss* pada setiap akhir *epoch* menggunakan metode *log loss*. Nilai metrik tersebut kemudian disuntikkan secara otomatis ke dalam log sistem pelatihan.

Guna menghindari masalah *overfitting* yang tersisa, penelitian ini menerapkan metode *Early Stopping* yang terintegrasi dengan pemantauan *validation loss*. Parameter kesabaran (*patience*) diatur sebesar 10 *epoch*. Nilai *patience* yang memadai ini memberikan model ruang toleransi untuk memulihkan performa sebelum akhirnya proses iterasi dihentikan secara otomatis. Konfigurasi penyimpanan bobot (*restore best weights*) dinonaktifkan, sehingga model mempertahankan status bobot terakhirnya tepat ketika kondisi penghentian otomatis terpicu.

Hasil riwayat iterasi dari keseluruhan proses pelatihan direkam dan divisualisasikan dalam bentuk grafik pergerakan nilai *loss* dan akurasi (*accuracy*) seperti yang disajikan pada Gambar 3. Grafik ini membandingkan penurunan nilai *Training Loss* dengan *Validation Loss*, serta peningkatan *Training Accuracy* dengan *Validation Accuracy* di setiap *epoch*, yang menjadi indikator evaluasi utama dalam membuktikan tingkat konvergensi, kestabilan, serta kemampuan generalisasi model terhadap data uji.



Gambar 3. Grafik Training and Validation Loss serta Accuracy

Berdasarkan visualisasi pada Gambar 3, terlihat bahwa pada awal fase pelatihan (hingga sekitar *epoch* ke-45), pergerakan metrik *loss* maupun akurasi cenderung stagnan. Nilai *loss* bertahan pada kisaran 0,65 dan akurasi tertahan pada kisaran 0,68. Kondisi awal ini wajar dan mengindikasikan bahwa model sedang mengeksplorasi ruang pencarian untuk menemukan gradien optimal guna keluar dari area minimum lokal (*local minima*). Setelah melewati fase eksplorasi tersebut, model mulai menunjukkan konvergensi yang sangat stabil, ditandai dengan kurva *loss* yang menurun dan kurva akurasi yang menanjak secara konsisten dan bersamaan, baik pada data latih maupun data validasi.

Keunggulan utama dari konfigurasi arsitektur ini terlihat sangat jelas ketika memasuki rentang *epoch* ke-100 hingga akhir iterasi. Berbeda dengan indikasi awal yang rentan terhadap penghafalan data, kurva pelatihan dan validasi pada grafik ini melandai secara berhimpitan tanpa memperlihatkan pemisahan jarak (*gap*) yang melebar. Pada akhir iterasi, *training loss* menyentuh angka 0,5467 dengan *training accuracy* sebesar 0,7493. Di saat yang sama, metrik validasi mengikutinya dengan sangat ketat, mencatatkan nilai *validation accuracy* sebesar 0,7472. Pergerakan kedua kurva yang sangat selaras ini merupakan bukti empiris yang kuat bahwa model terhindar sepenuhnya dari masalah *overfitting* dan memiliki tingkat generalisasi yang sangat ideal.

Meskipun performa model sudah sangat stabil, implementasi mekanisme *Early Stopping* tetap terbukti krusial dan bekerja secara presisi. Mekanisme ini menghentikan proses pelatihan secara otomatis pada *epoch* ke-139,

sebelum menyentuh batas maksimal 150 *epoch*. Penghentian ini dipicu karena sistem mendeteksi bahwa performa metrik validasi telah mencapai titik jenuh (*plateau*) optimalnya dan tidak lagi memperlihatkan perbaikan yang signifikan selama masa toleransi (*patience*). Dengan penghentian tersebut, model akhir yang dikembalikan adalah model pada titik konvergensi paling efisien, yang siap diimplementasikan untuk memberikan rekomendasi akurat kepada pengguna baru.

3.5 Evaluasi Performa Model

Setelah melalui proses pelatihan dan optimasi parameter, model DeepFM dievaluasi menggunakan data pengujian (*testing data*) yang tidak dilibatkan selama fase pelatihan. Evaluasi ini bertujuan untuk mengukur secara objektif sejauh mana sistem mampu memprediksi tingkat kecocokan serta mengurutkan rekomendasi program pemberdayaan secara presisi bagi pengguna baru. Pengukuran performa dilakukan menggunakan empat metrik utama yang sesuai untuk evaluasi klasifikasi dan *ranking*, yaitu AUC, *Precision@3*, *Recall@3*, dan *NDCG@3*. Hasil pengujian model dapat dilihat pada Tabel 4.

Tabel 4. Hasil Evaluasi Performa Model

Metrik	Nilai
AUC	0.7217
<i>Precision@3</i>	0.4144
<i>Recall@3</i>	0.8018
<i>NDCG@3</i>	0.7064

Tabel 4 menunjukkan bahwa arsitektur DeepFM yang diusulkan menghasilkan performa prediktif yang representatif dalam memodelkan interaksi antara profil minat RIASEC dan karakteristik program. Indikator keberhasilan pada kualitas urutan sistem rekomendasi ini ditunjukkan oleh perolehan nilai *Normalized Discounted Cumulative Gain* (*NDCG@3*) sebesar 0,7064. Nilai yang berhasil melampaui ambang batas 0,70 ini membuktikan bahwa sistem memiliki keandalan yang baik dalam menyusun peringkat (*ranking*) rekomendasi. Artinya, program pemberdayaan yang relevan dengan minat pemuda secara konsisten berhasil diprioritaskan dan ditempatkan pada posisi teratas dalam daftar rekomendasi yang dihasilkan.

Kemampuan model dalam mengklasifikasikan kesesuaian program juga terbukti stabil, ditunjukkan dengan nilai *Area Under the Curve* (AUC) sebesar 0,7217. Nilai ini mengindikasikan bahwa model memiliki kemampuan diskriminasi yang memadai dalam membedakan antara program yang benar-benar diminati dan yang tidak relevan bagi seorang pengguna. Lebih lanjut, keunggulan utama sistem ini ditunjukkan oleh perolehan metrik *Recall@3* sebesar 0,8018 (80,18%), yang mengonfirmasi sensitivitas model yang tinggi karena berhasil mengenali dan menangkap lebih dari 80% program relevan yang memang seharusnya direkomendasikan.

Di sisi lain, nilai *Precision@3* sebesar 0,4144 dinilai sudah optimal dan realistis. Angka ini merepresentasikan batasan matematis dari struktur perilaku responden pada *dataset*, di mana mayoritas pemuda secara alami hanya menaruh minat pada satu hingga dua pilihan program pemberdayaan dari total opsi yang tersedia. Keseluruhan hasil evaluasi ini menegaskan bahwa penerapan arsitektur DeepFM dengan konfigurasi *hyperparameter* yang optimal sukses menghasilkan model yang stabil dan terhindar dari bias kebocoran data (*data leakage*). Sistem yang dikembangkan terbukti valid secara empiris dan akurat dalam merekomendasikan program pemberdayaan yang terpersonalisasi guna mendukung pengembangan kapasitas pemuda.

3.6 Pembahasan

Berdasarkan hasil evaluasi eksperimen, penerapan model DeepFM terbukti efektif dalam menjembatani kesenjangan antara profil minat pemuda dan karakteristik program pemberdayaan yang tersedia. Capaian *Normalized Discounted Cumulative Gain* (*NDCG@3*) sebesar 0,7064 menunjukkan bahwa model memiliki keandalan yang baik dalam menyusun peringkat (*ranking*) rekomendasi, di mana program yang paling relevan secara konsisten berhasil diposisikan pada urutan prioritas tertinggi. Selain itu, nilai *Area Under the Curve* (AUC) sebesar 0,7217 mempertegas kemampuan diskriminasi model yang memadai dalam membedakan program yang relevan bagi pemuda. Performa ini didukung oleh *Recall@3* sebesar 0,8018 (80,18%) dan *Precision@3* sebesar 0,4144, yang mengonfirmasi bahwa integrasi komponen *Factorization Machine* (FM) dan *Deep Neural Network* (DNN) pada arsitektur DeepFM mampu secara optimal memproses data psikometrik yang bersifat kontinu dan padat (*dense*).

Untuk memvalidasi efektivitas arsitektur yang diusulkan, penelitian ini melakukan studi komparatif dengan melibatkan dua model pembandingan yaitu *Pure Deep Neural Network* (DNN) sebagai studi ablasi dan *XGBoost* (*Extreme Gradient Boosting*) sebagai representasi algoritma *machine learning* tradisional. Hasil pengujian komparatif disajikan pada Tabel 5.

Tabel 5. Hasil Uji Komparasi Model

Model Pembandingan	AUC	<i>Precision@3</i>	<i>Recall@3</i>	<i>NDCG@3</i>
--------------------	-----	--------------------	-----------------	---------------

<i>Pure DNN</i>	0.6679	0.4099	0.7725	0.6577
XGBoost	0.7250	0.4144	0.7860	0.7017
DeepFM	0.7217	0.4144	0.8018	0.7064

Berdasarkan Tabel 5, model *Pure DNN* mencatatkan performa terendah, mengonfirmasi kelemahan arsitektur saraf tiruan murni dalam menangkap korelasi antar fitur secara eksplisit tanpa bantuan *Factorization Machine*. Sementara itu, XGBoost menunjukkan performa kompetitif dengan nilai AUC 0,7250. Namun, dalam konteks sistem rekomendasi, metrik pemeringkatan (NDCG) dan sensitivitas *Recall* lebih esensial. Pada indikator tersebut, DeepFM berhasil mengungguli XGBoost, membuktikan bahwa integrasi komponen FM pada DeepFM bekerja lebih optimal dalam memetakan interaksi laten antara profil minat RIASEC dengan karakteristik program.

Jika dikomparasikan dengan penelitian terdahulu, hasil penelitian ini menunjukkan posisi yang sangat kompetitif. Wen dan Geng [8] memperoleh nilai AUC 0,9341 pada domain pendidikan anak usia dini. Meskipun nilai AUC penelitian ini (0,7217) lebih rendah, perbedaan ini bersifat rasional karena penelitian ini menggunakan data psikometrik statis tanpa riwayat interaksi pengguna yang masif, namun tetap menunjukkan efisiensi konvergensi yang konsisten.

Selanjutnya, jika dibandingkan dengan penelitian dari Hidayat dkk. [9] yang menggunakan model *Hybrid Neural Network* dan *Collaborative Filtering* untuk prediksi kesesuaian program studi (menghasilkan *precision* 87% dan *recall* 82%), model DeepFM pada penelitian ini mencatatkan nilai *recall@3* sebesar 80,18%. Meskipun capaian *recall* pada penelitian ini berada sedikit di bawah penelitian tersebut, nilai 80,18% tetap menunjukkan performa yang sangat kompetitif dan berada pada level performa yang sebanding. Perbedaan tipis ini dipengaruhi oleh perbedaan cakupan domain; penelitian Hidayat dkk. [9] berfokus pada prediksi pilihan program studi yang cenderung bersifat formal, terstruktur, dan memiliki kriteria kecocokan yang sangat spesifik. Sebaliknya, karakteristik program kepemudaan yang dikelola Kemenpora memiliki spektrum yang jauh lebih luas dan fleksibel.

Terkait dengan nilai *precision@3* sebesar 41,44%, perbedaan yang signifikan jika dibandingkan dengan penelitian terdahulu [9] adalah konsekuensi logis dari fleksibilitas domain ini. Dalam konteks program kepemudaan, satu individu sering kali memiliki kecocokan yang valid untuk berbagai tipe program sekaligus, sehingga daftar rekomendasi yang dihasilkan memiliki variasi yang lebih banyak dibandingkan prediksi program studi formal yang kaku. Dengan demikian, model DeepFM dalam penelitian ini tetap terbukti berhasil menjalankan perannya secara optimal untuk memberikan rekomendasi yang relevan, kaya akan opsi, dan adaptif terhadap profil minat pemuda yang multidimensi.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, arsitektur DeepFM terbukti mampu mengintegrasikan profil minat RIASEC dengan karakteristik program pemberdayaan pemuda secara efektif. Melalui pengujian empiris, model DeepFM berhasil mencatatkan nilai AUC sebesar 0,7217, *recall@3* sebesar 0,8018, *precision@3* sebesar 0,4144 dan *NDCG@3* sebesar 0,7064. Hasil komparasi menunjukkan bahwa DeepFM memiliki keunggulan fungsional dalam aspek pemeringkatan dan sensitivitas penangkapan program relevan dibandingkan dengan model *Pure DNN* maupun XGBoost. Hal ini membuktikan bahwa sinergi antara komponen *Factorization Machine* dan *Deep Neural Network* pada arsitektur DeepFM memberikan representasi fitur yang lebih optimal dalam memproses data psikometrik multidimensi yang bersifat kontinu.

Meskipun performa model telah terbukti kompetitif, penelitian ini masih memiliki keterbatasan pada ukuran *dataset* yang berskala moderat dengan 488 responden, sehingga generalisasi model terhadap variasi populasi pemuda yang lebih luas masih perlu dikembangkan. Selain itu, fitur masukan yang digunakan saat ini masih bersifat statis dan belum mengakomodasi variabel dinamis seperti rekam jejak pelatihan atau perkembangan keahlian pengguna seiring berjalannya waktu. Keterbatasan ini memberikan celah bagi pengembangan sistem yang lebih adaptif di masa mendatang.

Oleh karena itu, untuk pengembangan penelitian selanjutnya, disarankan agar memperluas volume data dan keberagaman program guna menguji skalabilitas model pada skala yang lebih besar. Selain itu, penerapan algoritma tingkat lanjut untuk penanganan ketidakseimbangan kelas, seperti SMOTE atau penerapan pembobotan fungsi kerugian (*focal loss*), dapat dieksplorasi lebih jauh untuk meminimalkan bias prediksi pada program minoritas, sangat disarankan untuk meningkatkan presisi prediksi pada kategori program dengan partisipasi rendah. Dari sisi arsitektur fitur, penelitian selanjutnya diharapkan mampu mengintegrasikan representasi pengguna yang lebih kaya dan dinamis. Pemanfaatan data historis berupa riwayat aktivitas secara sekuensial atau parameter latar belakang keahlian yang lebih spesifik diyakini akan mampu memandu model dalam menghasilkan sistem personalisasi rekomendasi yang jauh lebih adaptif, komprehensif, dan tepat sasaran.

REFERENCES

- [1] E. Wihardjo dkk., "Peluang Dan Tantangan Indonesia Emas 2045," 2025, hlm. 1–226. Diakses: 23 Mei 2026. [Daring]. Tersedia pada: <https://penerbitlitnus.co.id/portfolio/peluang-dan-tantangan-indonesia-emas-2045/>

- [2] Kemenpora, “Draf Rencana Strategis Kementerian Pemuda dan Olahraga Tahun 2025 - 2029,” 2025.
- [3] A. P. Sudarso, F. K. Qaulibi, M. Maira, dan M. R. F, “Pengabdian Masyarakat Pemberdayaan Melalui Program ‘Teman’ (Tempa Mental Dan Minat): Strategi Peningkatan Karakter Minat dan Kepercayaan Diri Siswa SMK Buana Mulia Parung Kabupaten Bogor,” *J. Abdi Raya Nusant.*, vol. 1, no. 2, hlm. 177–189, Okt 2025, doi: 10.70285/ffw56w57.
- [4] B. Triatmanto, G. Apriyanto, dan S. H. dkk, *Model Pemberdayaan Masyarakat Holistik : Berorientasi Potensi Lokal*. Uwais Inspirasi Indonesia, 2024.
- [5] J. L. Holland, *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments*. Prentice-Hall, 1985.
- [6] H. Guo, R. Tang, Y. Ye, Z. Li, dan X. He, “DeepFM: A Factorization-Machine based Neural Network for CTR Prediction,” hlm. 1725–1731, 2017.
- [7] Z. R. R. R. Faqih, F. Metandi, dan Hadiwijaya, “Implementasi Sistem Rekomendasi Game Menggunakan Hybrid Filtering Pada Platfrom Digital Steam,” no. Vol. 3 No. 6 (2025): JURNAL MEDIA AKADEMIK Edisi Juni, Jun 2025, doi: <https://doi.org/10.62281/v3i6.2363>.
- [8] L. Wen dan J. Geng, “Research on Personalized Early Childhood Education Recommendation Model Based on Improved DeepFM,” dalam *Proceedings of the 2023 4th International Conference on Computer Science and Management Technology*, dalam ICCSMT '23. New York, NY, USA: Association for Computing Machinery, Apr 2024, hlm. 993–996. doi: 10.1145/3644523.3644702.
- [9] S. Z. Hidayat, Y. V. Via, dan H. Maualana, “Sistem Rekomendasi Program Studi Jenjang Sarjana Untuk Siswa Smk Berbasis Self-Assessment Riasec Dengan Metode Hybrid Neural Network Dan Collaborative Filtering,” *JATI J. Mhs. Tek. Inform.*, vol. 9, no. 3, hlm. 4318–4326, Mei 2025, doi: 10.36040/jati.v9i3.13642.
- [10] A. L. P. Sulistyio dkk., “Analisis Teori Karier Holland dalam Mendukung Bimbingan dan Konseling Karier,” *Pros. Konseling Kearifan Nusant. KKN*, vol. 4, hlm. 901–909, Jan 2025, doi: 10.29407/14x7b281.
- [11] Y. Zhang, “An Introduction to Matrix factorization and Factorization Machines in Recommendation System, and Beyond,” 12 Maret 2022, *arXiv: arXiv:2203.11026*. doi: 10.48550/arXiv.2203.11026.
- [12] J. Lu, Y. Long, X. Li, Y. Shen, dan X. Wang, “Hybrid Model Integration of LightGBM, DeepFM, and DIN for Enhanced Purchase Prediction on the Elo Dataset,” dalam *2024 IEEE 7th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Sep 2024, hlm. 16–20. doi: 10.1109/ICISCAE62304.2024.10761856.
- [13] Ardiansyah, Risnita, dan M. S. Jailani, “Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif,” *IHSAN J. Pendidik. Islam*, vol. 1, no. 2, hlm. 1–9, Jul 2023, doi: 10.61104/ihsan.v1i2.57.
- [14] Kemenpora, “Pemberitahuan Tertulis Program Kegiatan Kepemudaan.” 1 Desember 2025.
- [15] N. W. A. Ulandari, I. P. W. Putra, N. M. Astiti, N. L. Hermawan, dan P. G. G. Putra, “Implementasi Holland Riasec Model Untuk Rekomendasi Program Studi Calon Mahasiswa Baru Itb Stikom Bali: indonesia,” *J. Pendidik. Teknol. Dan Kejur.*, vol. 23, no. 1, hlm. 93–104, Jan 2026, doi: 10.23887/jptk-undiksha.v23i1.104327.
- [16] M. Ma, G. Wang, dan T. Fan, “Improved DeepFM Recommendation Algorithm Incorporating Deep Feature Extraction,” *Appl. Sci.*, vol. 12, no. 23, hlm. 11992, Jan 2022, doi: 10.3390/app122311992.
- [17] F. Arasadina, S. Briandoko, dan M. A. Setiawan, “Prediksi Pembatalan Pemesanan Hotel: Random Forest dan XGBoost dengan Pipeline Leak-Free dan Ablation Study: Hotel Booking Cancellation Prediction: Random Forest and XGBoost with Leak-Free Pipeline and Ablation Study,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 6, no. 3, hlm. 1018–1030, Jun 2026, doi: 10.57152/malcom.v6i3.2674.
- [18] R. A. Pratama, Y. Safi'i, M. A. Nugraha, A. S. Sobihah, dan N. Ifada, “Perbandingan User-Based dan Item-Based pada Sistem Rekomendasi Film Kombinasi Teknik Reduksi Dimensi dan Clustering,” *J. Tekno Insentif*, vol. 19, no. 1, hlm. 1–14, Jul 2025, doi: 10.36787/jti.v19i1.1662.
- [19] Ermanto dan N. Surojudin, “Perbandingan Metode Klasifikasi dalam Memprediksi Penyakit Ginjal Kronis,” *J. Inform. Ekon. Bisnis*, hlm. 715–723, Sep 2025, doi: 10.37034/infv7i3.1263.
- [20] A. Hermawan, A. Kanugraha, I. F. Afgani, K. N. Safaati, dan M. A. A. Ramadhani, “Sistem Rekomendasi Musik Spotify Berbasis Pendekatan Hybrid Alternating Least Square Dan Content-Based Filtering,” *Modem J. Inform. Dan Sains Teknol.*, vol. 4, no. 2, hlm. 43–60, Apr 2026, doi: 10.62951/modem.v4i2.869.
- [21] S. Hendrawan, “Optimasi Sistem Rekomendasi Film Menggunakan Metode Hybrid Filtering,” skripsi, KODEUNIVERSITAS041060#UniversitasBuddhiDharma, 2025. Diakses: 4 Juni 2026. [Daring]. Tersedia pada: <http://repository.buddhidharma.ac.id>
- [22] D. A. Nugroho, C. Lubis, dan N. J. Perdana, “Sistem Rekomendasi Film Menggunakan Metode Neural Collaborative Filtering,” *J. Inf. Technol. Comput. Sci.*, Jun 2024, doi: <https://doi.org/10.31539/intecom.v7i3.8033>.
- [23] A. A. Saputra, B. N. Sari, dan C. Rozikin, “Penerapan Algoritma Extreme Gradient Boosting (Xgboost) Untuk Analisis Risiko Kredit,” *J. Ilm. Wahana Pendidik.*, vol. 10, no. 7, hlm. 27–36, Apr 2024, doi: 10.5281/zenodo.10960080.



- [24] A. K. E. Pily, Susanti, U. Rio, dan Tashid, “Komparasi K-Means Clustering dengan Euclidean dan Cosine Similarity untuk Segmentasi dan Rekomendasi Produk pada Data E-Commerce,” *Indones. J. Comput. Sci.*, vol. 14, no. 2, Apr 2025, doi: 10.33022/ijcs.v14i2.4713.
- [25] J. Rounds, C. M. Walker, S. X. Day, L. Hubert, P. Lewis, dan D. Rivkin, “O*NET Interest Profiler: Reliability, Validity, and Self-Scoring,” National Center for O*NET Development, Raleigh, North Carolina 27611, 1999. Diakses: 25 Mei 2026. [Daring]. Tersedia pada: https://www.onetcenter.org/dl_files/IP_RVS.pdf
- [26] R. M. Putri, Y. Yosef, F. Rozzaqyah, M. Vyanti, A. Nadya, dan M. Nisa, “Validitas Dan Reliabilitas I.Riasec (Inventori Karir Berbasis Teori Holland Dalam Konteks Pendidikan Indonesia,” *Al-Isyraq J. Bimbing. Penyul. Dan Konseling Islam*, vol. 7, no. 3, hlm. 847–864, Des 2024, doi: 10.59027/alisyraq.v7i3.808.
- [27] Resti Kurnia Triastanti dan Arini Hardianti, “Uji Validitas dan Reliabilitas Kuesioner Tipe Chronotype pada Remaja,” *INSOLOGI J. Sains Dan Teknol.*, vol. 3, no. 1, hlm. 118–124, Feb 2024, doi: 10.55123/insologi.v3i1.3172.