

**PENERAPAN METODE *SENTENCE-BERT* UNTUK SISTEM  
REKOMENDASI PEMILIHAN PROGRAM STUDI  
SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi  
Sistem Informasi



disusun oleh

**RIFANDI SAPTO AJI**

**22.12.2455**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2026**

**PENERAPAN METODE *SENTENCE-BERT* UNTUK SISTEM  
REKOMENDASI PEMILIHAN PROGRAM STUDI**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi

Sistem Informasi



disusun oleh

**RIFANDI SAPTO AJI**

**22.12.2455**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**PENERAPAN METODE *SENTENCE-BERT* UNTUK SISTEM  
REKOMENDASI PEMILIHAN PROGRAM STUDI**

yang disusun dan diajukan oleh

**Rifandi Sapto Aji**

**22.12.2455**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 25 Mei 2026

**Dosen Pembimbing,**



**Ir. Rum Mohamad Andri K Rasyid, M.Kom.**

**NIK. 190302011**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**PENERAPAN METODE SENTENCE-BERT UNTUK SISTEM  
REKOMENDASI PEMILIHAN PROGRAM STUDI**

yang disusun dan diajukan oleh

**Rifandi Sapto Aji**

**22.12.2455**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 25 Mei 2026

**Susunan Dewan Penguji**

**Nama Penguji**

Irma Rofni Wulandari, S.pd., M.Eng.  
**NIK. 190302329**

Ikamah, S.Kom., M.Kom.  
**NIK. 190302282**

Ir. Rum Mohamad Andri K Rasyd, M.Kom.  
**NIK. 190302011**

**Tanda Tangan**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 25 Mei 2026

**DEKAN FAKULTAS ILMU KOMPUTER**



Prof. Dr. Kusrini, S.Kom., M.Kom.  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Rifandi Sapto Aji  
NIM : 22.12.2455

Menyatakan bahwa Skripsi dengan judul berikut:

### **PENERAPAN METODE SENTENCE-BERT UNTUK SISTEM REKOMENDASI PEMILIHAN PROGRAM STUDI**

Dosen Pembimbing : Ir. Rum Mohamad Andri K Rasyid, M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 25 Mei 2026

Yang Menyatakan,

  
Rifandi Sapto Aji

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan lancar. Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan Program Sarjana pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Selain itu penulis dengan segala kerendahan hati ingin menyampaikan rasa terima kasih kepada semua pihak yang telah berjasa memberikan dukungan dan bantuan untuk menyelesaikan laporan tugas akhir ini. Untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta
2. Ibu Prof. Dr. Kusrini, S.Kom., M.Kom., selaku Dekan Program Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartanto, M.Kom., selaku Ketua Program Studi Sistem Informasi
4. Ir. Rum Mohamad Andri K Rasyid, M.Kom. selaku dosen pembimbing yang memberikan arahan, saran, dan motivasi terhadap penulis
5. Kedua orang tua, keluarga besar, dan teman-teman tercinta yang memberikan semangat dan doa kepada penulis.
6. FC Barcelona yang telah menjadi tontonan saya selagi bergadang mengerjakan skripsi ini.

Yogyakarta, 25 Mei 2026

Rifandi Sapto Aji

## DAFTAR ISI

|                                                      |     |
|------------------------------------------------------|-----|
| HALAMAN JUDUL .....                                  | i   |
| HALAMAN PERSETUJUAN.....                             | ii  |
| HALAMAN PENGESAHAN .....                             | iii |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI .....            | iv  |
| KATA PENGANTAR .....                                 | v   |
| DAFTAR ISI.....                                      | vi  |
| DAFTAR TABEL.....                                    | ix  |
| DAFTAR GAMBAR.....                                   | x   |
| DAFTAR LAMBANG DAN SINGKATAN .....                   | xi  |
| DAFTAR ISTILAH .....                                 | xii |
| INTISARI .....                                       | xiv |
| ABSTRACT.....                                        | xv  |
| BAB I PENDAHULUAN.....                               | 1   |
| 1.1 Latar Belakang .....                             | 1   |
| 1.2 Rumusan Masalah.....                             | 2   |
| 1.3 Batasan Masalah .....                            | 3   |
| 1.4 Tujuan Penelitian .....                          | 3   |
| 1.5 Manfaat Penelitian .....                         | 3   |
| 1.6 Sistematika Penulisan .....                      | 4   |
| BAB II TINJAUAN PUSTAKA .....                        | 5   |
| 2.1 Studi Literatur .....                            | 5   |
| 2.2 Dasar Teori .....                                | 9   |
| 2.3 Sistem Rekomendasi.....                          | 9   |
| 2.4 Metode Dalam Sistem Rekomendasi .....            | 10  |
| 2.4.1 Collaborative Filtering .....                  | 10  |
| 2.4.2 Content-Based Filtering .....                  | 11  |
| 2.4.3 Pendekatan Berbasis Representasi Semantik..... | 12  |
| 2.5 Text Processing (Pra-pemrosesan Teks).....       | 12  |

|                                   |                                                           |    |
|-----------------------------------|-----------------------------------------------------------|----|
| 2.6                               | Sentence-BERT .....                                       | 13 |
| 2.7                               | Cosine similarity .....                                   | 14 |
| 2.8                               | Evaluasi Sistem Rekomendasi .....                         | 15 |
| 2.8.1                             | Precision.....                                            | 16 |
| 2.8.2                             | Analisis Semantik .....                                   | 16 |
| 2.8.3                             | Mean Reciprocal Rank (MRR) .....                          | 17 |
| BAB III METODE PENELITIAN .....   |                                                           | 18 |
| 3.1                               | Objek Penelitian.....                                     | 18 |
| 3.2                               | Alur Penelitian .....                                     | 18 |
| 3.2.1                             | Pengumpulan Data .....                                    | 19 |
| 3.2.2                             | Preprocessing Data.....                                   | 20 |
| 3.2.3                             | Representasi Teks Menggunakan Sentence-BERT .....         | 20 |
| 3.2.4                             | Perhitungan Kemiripan Menggunakan Cosine Similarity ..... | 21 |
| 3.2.5                             | Rekomendasi Program Studi.....                            | 23 |
| 3.2.6                             | Evaluasi Sistem.....                                      | 23 |
| 3.3                               | Alat dan Bahan.....                                       | 24 |
| 3.3.1                             | Perangkat Keras (Hardware).....                           | 24 |
| 3.3.2                             | Perangkat Lunak (Software) .....                          | 25 |
| 3.3.3                             | Bahan Penelitian .....                                    | 25 |
| BAB IV HASIL DAN PEMBAHASAN ..... |                                                           | 27 |
| 4.1                               | Pengumpulan Data .....                                    | 27 |
| 4.2                               | Pra-pemrosesan Teks (Teks Pre-processing).....            | 29 |
| 4.2.1                             | Case Folding .....                                        | 30 |
| 4.2.2                             | Tokenization .....                                        | 31 |
| 4.2.3                             | Stopword Removal.....                                     | 32 |
| 4.2.4                             | Stemming .....                                            | 33 |
| 4.3                               | Ekstraksi Fitur Menggunakan Sentence-BERT (SBERT).....    | 34 |

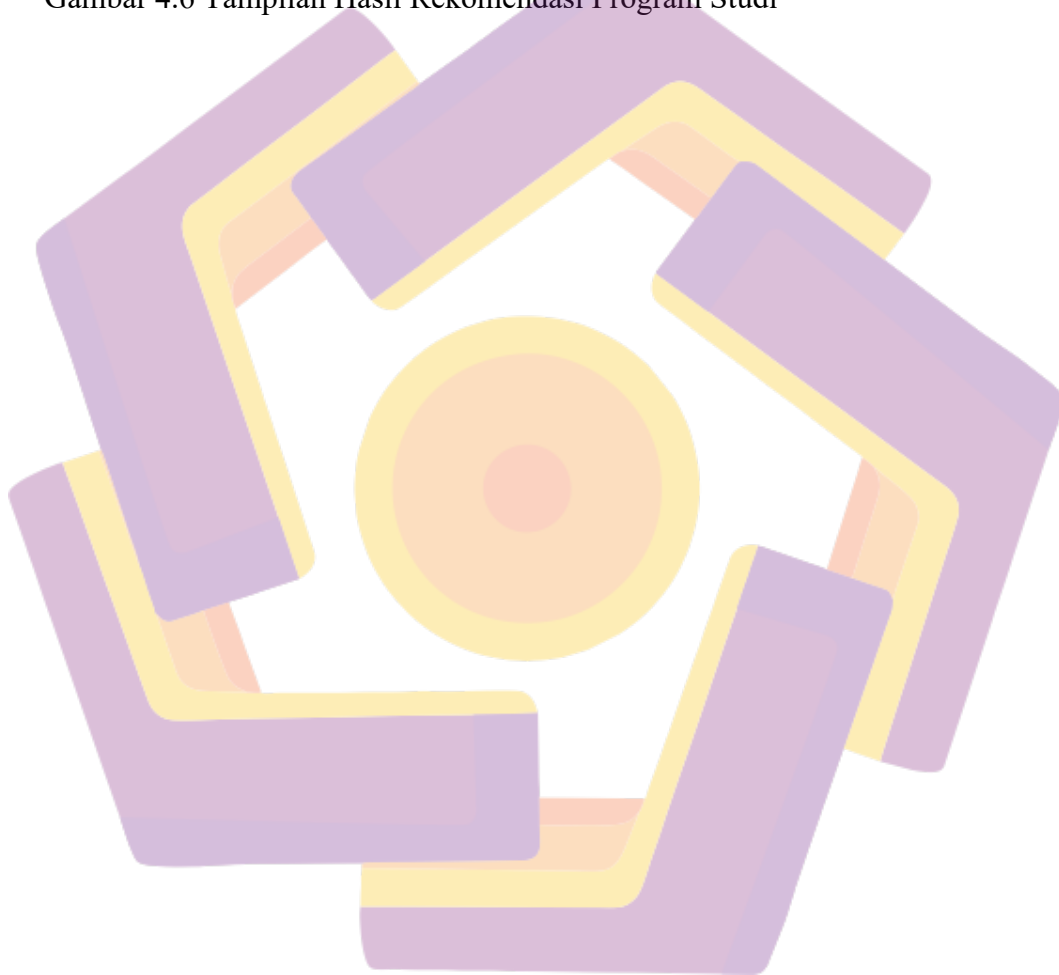
|                                                             |    |
|-------------------------------------------------------------|----|
| 4.4 Cosine Similarity .....                                 | 36 |
| 4.4.1 Penentuan Sampel Vektor dan Asal Ekstraksi Data ..... | 37 |
| 4.4.2 Proses Perhitungan Matematis .....                    | 37 |
| 4.4.3 Hasil Evaluasi Kemiripan .....                        | 39 |
| 4.5 Hasil Rekomendasi .....                                 | 40 |
| 4.6 Evaluasi.....                                           | 41 |
| 4.6.1 Analisis Semantik .....                               | 42 |
| 4.6.2 Precision @5 .....                                    | 42 |
| 4.6.3 Mean Reciprocal Rank (MRR) .....                      | 43 |
| 4.7 Implementasi Antarmuka.....                             | 44 |
| BAB V PENUTUP .....                                         | 47 |
| 5.1 Kesimpulan .....                                        | 47 |
| 5.2 Saran .....                                             | 48 |
| REFERENSI .....                                             | 49 |

## DAFTAR TABEL

|                                                    |    |
|----------------------------------------------------|----|
| Tabel 2.1 Keaslian Penelitian                      | 4  |
| Tabel 4.1 Atribut Data                             | 22 |
| Tabel 4.2 Sampel Detail Entitas Data Program Studi | 23 |
| Tabel 4.3 Hasil Case Folding                       | 25 |
| Tabel 4.4 Hasil Tokenization                       | 26 |
| Tabel 4.5 Hasil Stopword Removal                   | 27 |
| Tabel 4.6 Hasil Stemming                           | 28 |
| Tabel 4.7 Hasil Sentence-BERT (SBERT)              | 29 |
| Tabel 4.8 Perhitungan Cosine Similarity            | 31 |
| Tabel 4.9 Hasil Rekomendasi                        | 34 |
| Tabel 4.10 Hasil Analisis Semantik                 | 36 |
| Tabel 4.11 Hasil Precision@5                       | 37 |
| Tabel 4.12 Hasil Mean Reciprocal Rank (MRR)        | 38 |

## DAFTAR GAMBAR

|                                                               |    |
|---------------------------------------------------------------|----|
| Gambar 3.1 Alur Penelitian                                    | 13 |
| Gambar 4.2 Alur Pra-pemrosesan                                | 24 |
| Gambar 4.3 Tangkapan Layar Eksekusi Program pada Google Colab | 35 |
| Gambar 4.5 Tampilan Antarmuka Input Sistem                    | 40 |
| Gambar 4.6 Tampilan Hasil Rekomendasi Program Studi           | 41 |



## DAFTAR LAMBANG DAN SINGKATAN

|         |                                   |
|---------|-----------------------------------|
| SBERT   | Sentence-BERT                     |
| LLDIKTI | Lembaga Layanan Pendidikan Tinggi |
| PTS     | Perguruan Tinggi Swasta           |
| NLP     | Natural Language Processing       |
| MRR     | Mean Reciprocal Rank              |
| RMSE    | Root Mean Squared Error           |
| MAE     | Mean Absolute Error               |
| AI      | Artificial Intelligence           |
| UI/UX   | User Interface / User Experience  |

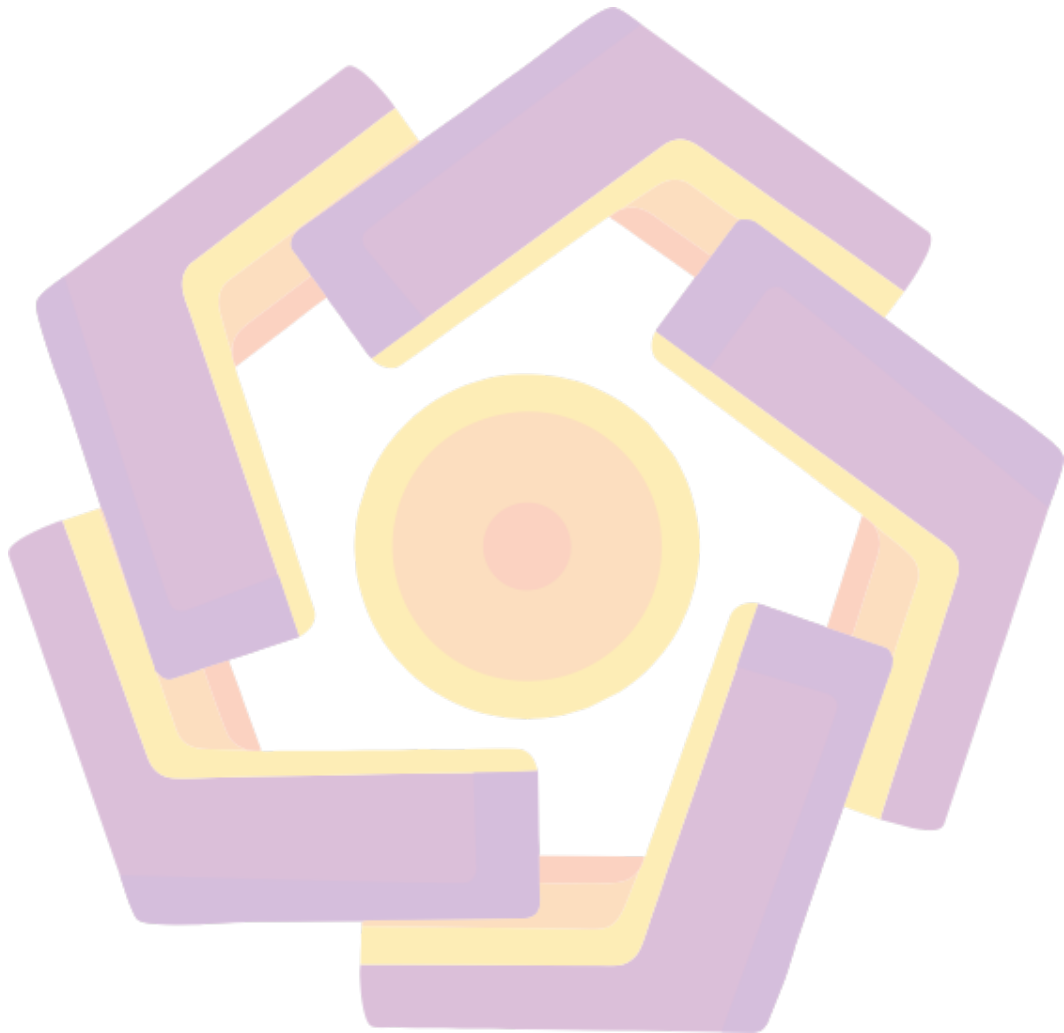


## DAFTAR ISTILAH

|                                  |                                                                                                                                                      |
|----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sentence-BERT (SBERT)            | Pengembangan dari model BERT yang dirancang untuk menghasilkan representasi vektor pada tingkat kalimat guna memahami konteks makna secara mendalam  |
| Cosine Similarity                | Metode untuk mengukur tingkat kesamaan antara dua teks dengan membandingkan arah vektor representasinya dalam ruang vektor                           |
| Text Processing (Pra-pemrosesan) | Tahap awal pengolahan data teks untuk mengubah teks mentah menjadi bentuk yang lebih terstruktur                                                     |
| Case Folding                     | Tahap mengubah seluruh huruf menjadi huruf kecil serta menghapus angka dan tanda baca untuk menstandarisasi data                                     |
| Tokenization                     | Proses memecah untaian teks atau kalimat menjadi potongan-potongan kata tunggal yang disebut token                                                   |
| Stopword Removal                 | Proses eliminasi kata-kata umum (seperti kata hubung atau kata depan) yang tidak memiliki bobot fitur penting dalam analisis teks                    |
| Stemming                         | Proses mengembalikan kata-kata yang memiliki imbuhan morfologis menjadi bentuk kata dasarnya (root word)                                             |
| Word Embedding                   | Representasi numerik matematis yang mentransformasikan teks bahasa manusia menjadi angka agar dapat diproses oleh mesin                              |
| Mean Pooling                     | Operasi komputasi untuk menghitung nilai rata-rata dari seluruh vektor kata guna menghasilkan satu representasi vektor tunggal bagi keseluruhan teks |
| Semantic Search                  | Pendekatan pencarian yang didasarkan pada kedekatan makna antar konsep, bukan sekadar kesamaan kata kunci secara harafiah                            |
| Precision @5                     | Metrik evaluasi yang mengukur proporsi item relevan dalam lima daftar rekomendasi teratas yang diberikan oleh sistem                                 |

Mean Reciprocal Rank (MRR)

Metrik evaluasi yang berfokus pada posisi peringkat di mana rekomendasi relevan pertama kali muncul dalam daftar hasil



## INTISARI

Pemilihan program studi merupakan keputusan penting bagi calon mahasiswa agar sesuai dengan minat dan arah pengembangan karier di masa depan. Namun, banyaknya pilihan program studi pada bidang komputer sering kali menimbulkan kebingungan karena informasi yang tersedia masih berupa teks deskriptif yang belum diolah secara terstruktur untuk mendukung pengambilan keputusan. Oleh karena itu, penelitian ini bertujuan untuk merancang dan membangun sistem rekomendasi pemilihan program studi berbasis web menggunakan metode *Sentence-BERT* (SBERT). Penelitian ini menggunakan dataset 23 deskripsi program studi yang diekstraksi dari 8 Perguruan Tinggi Swasta (PTS) di bawah naungan LLDIKTI Wilayah V Daerah Istimewa Yogyakarta. Tahapan penelitian meliputi pra-pemrosesan teks (*case folding*, *tokenization*, *stopword removal*, dan *stemming*), ekstraksi fitur menggunakan arsitektur SBERT (model *paraphrase-multilingual-MiniLM-L12-v2*), perhitungan tingkat kemiripan semantik menggunakan *Cosine Similarity*, serta implementasi antarmuka menggunakan *framework* Streamlit. Hasil pengujian menunjukkan bahwa model SBERT berhasil memahami konteks dan kedekatan makna semantik antara input minat pengguna dengan kurikulum program studi, bukan sekadar melakukan pencocokan kata kunci secara leksikal. Evaluasi sistem membuktikan efektivitas yang sangat baik, dibuktikan dengan capaian nilai *Precision@5* sebesar 60% dan skor sempurna pada *Mean Reciprocal Rank* (MRR) sebesar 1.0. Sistem rekomendasi ini diharapkan dapat membantu calon mahasiswa dalam mengambil keputusan pemilihan program studi secara lebih objektif dan relevan berdasarkan penjabaran minat mereka.

**Kata kunci:** sistem rekomendasi, pemilihan program studi, perguruan tinggi, bidang komputer, sistem informasi.

## ABSTRACT

*Choosing an appropriate study program is a crucial decision for prospective students to align with their interests and future career development. However, the abundance of computer-related study programs often causes confusion because the available information is generally presented in the form of unstructured descriptive text. Therefore, this research aims to design and develop a web-based recommendation system for selecting study programs using the Sentence-BERT (SBERT) method. This study utilized a dataset of 23 study program descriptions extracted from 8 private higher education institutions (PTS) under the auspices of LLDIKTI Region V, Special Region of Yogyakarta. The research stages included text pre-processing (case folding, tokenization, stopwords removal, and stemming), feature extraction using the SBERT architecture (paraphrase-multilingual-MiniLM-L12-v2 model), semantic similarity calculation using Cosine Similarity, and user interface implementation using the Streamlit framework. The test results showed that the SBERT model successfully understood the semantic context and meaning closeness between user interests and study program curricula, rather than simply relying on literal keyword matching. The system evaluation demonstrated excellent effectiveness, evidenced by a Precision@5 score of 60% and a perfect Mean Reciprocal Rank (MRR) score of 1.0. This recommendation system is expected to assist prospective students in making more objective and relevant study program selection decisions based on their interest descriptions.*

**Keyword:** *recommendation system, study program selection, higher education, computer field, information system.*