

## **Aplikasi Sistem Prediksi Mahasiswa Penerima Beasiswa Berbasis Web dengan Menerapkan Model Klasifikasi *K-Nearest Neighbors***

**Dede Kurniadi<sup>1\*</sup>, Fitri Nuraeni<sup>2</sup>, Aura Fitria Hazar<sup>3</sup>**  
<sup>1,2,3</sup>Institut Teknologi Garut, Indonesia

\*email: dede.kurniadi@itg.ac.id

---

### **Info Artikel**

Dikirim: 14 Agustus 2023  
Diterima: 5 September 2023  
Diterbitkan: 18 Mei 2024

---

### **Kata kunci:**

Beasiswa;  
*K-Nearest Neighbors*;  
Sistem Prediksi;  
*Waterfall*.

---

### **ABSTRAK**

Beasiswa Kartu Indonesia Pintar Kuliah atau KIP-K merupakan salah satu dari banyaknya beasiswa yang disediakan oleh pemerintah untuk melanjutkan pendidikan ke jenjang tinggi bagi para pelajar yang berprestasi namun terkendala oleh biaya. Salah satu perguruan tinggi di Garut yang menyediakan penerimaan mahasiswa baru melalui jalur beasiswa ini adalah Institut Teknologi Garut. Setiap tahunnya, Institut Teknologi Garut selalu mengalami kenaikan jumlah pendaftar beasiswa KIP-K akan tetapi hal ini tidak sebanding dengan jumlah kuota yang didapatkan sehingga harus dilakukan proses seleksi agar beasiswa dapat tepat sasaran. Proses seleksi sendiri dilakukan secara manual tanpa bantuan sistem khusus yang dapat membantu menyeleksi dengan lebih tepat dan efisien. Tujuan penelitian ini untuk membangun aplikasi sistem prediksi berbasis web dengan menerapkan model klasifikasi *K-Nearest Neighbors* dalam membantu menyeleksi calon penerima beasiswa KIP-K di Institut Teknologi Garut berdasarkan skor hasil tes, kondisi ekonomi, prestasi akademik dan non akademik dari setiap peserta. Model klasifikasi diterapkan dalam sistem sebagai proses pengklasifikasian kelayakan calon penerima agar proses seleksi lebih terfokus pada peserta yang terkategori layak. Sistem dibangun dengan menggunakan metode pendekatan *waterfall* agar pembangunan sistem lebih terstruktur. Penelitian ini menghasilkan aplikasi berupa sistem prediksi berbasis web yang dapat membantu mengklasifikasikan kelayakan serta menyeleksi calon penerima beasiswa KIP-K di Institut Teknologi Garut dengan tingkat akurasi sistem dalam memprediksi kelayakan peserta sebesar 91,86%.

---

## **1. PENDAHULUAN**

Beasiswa adalah pemberian bantuan keuangan kepada individu tertentu yang bertujuan agar dapat digunakan demi keberlangsungan pendidikan yang ditempuh [1]. Terdapat banyak beasiswa yang disediakan oleh pemerintah, salah satunya PIP (Program Indonesia Pintar) untuk perguruan tinggi diberikan dalam bentuk KIP-K (Kartu Indonesia Pintar Kuliah) [2]. Institut Teknologi Garut merupakan salah satu perguruan tinggi di Garut yang setiap tahunnya membuka penerimaan mahasiswa baru melalui jalur beasiswa ini. Berdasarkan hasil penelitian [3] jumlah pendaftar beasiswa KIP-K yang ditolak di Institut Teknologi Garut pada tahun 2016-2018 hampir sebanding dengan kuota penerimanya, hal tersebut terjadi karena jumlah pendaftar yang semakin naik namun kuota penerimaan yang diberikan oleh pemerintah terbatas dan hanya naik sedikit setiap tahunnya. Sehingga harus dilakukan penyeleksian yang ketat oleh bagian kemahasiswaan agar penyaluran beasiswa dapat tepat sasaran [4]. Selain itu, berdasarkan hasil wawancara proses penyeleksian calon mahasiswa penerima beasiswa dilakukan secara manual tanpa adanya bantuan dari sistem khusus yang mampu untuk memprediksi

kelayakan calon penerima beasiswa hanya dengan memasukan seluruh data pendaftar dan akan langsung muncul hasil klasifikasinya.

Berikut beberapa penelitian rujukan yang digunakan, yaitu penelitian dari Kurniadi, dkk [3] mengenai penerapan algoritma *K-Nearest Neighbors* dan Teknik *SMOTE* dalam memprediksi kelayakan calon mahasiswa penerima bantuan beasiswa KIP-K di Institut Teknologi Garut dengan menghasilkan model klasifikasi dengan tingkat akurasi model sebesar 90,12%. Penelitian oleh Mulyani, dkk [4] mengenai prediksi penerima beasiswa PPA dan KIP-K dengan menggunakan algoritma *Naïve Bayes* menggunakan 100 data untuk KIP-K dan 45 data PPA yang menghasilkan model klasifikasi dengan akurasi 80% untuk PPA dan 91% KIP-K. Kemudian Haryanto, dkk [5] meneliti mengenai perancangan sistem pendukung keputusan yang dapat menentukan penerima beasiswa di Institut Teknologi Telkom Purwokerto dengan menggunakan metode *Artificial Neural Network* dengan menggunakan metode pengembangan sistem *Waterfall*. Selain itu penelitian mengenai sistem pendukung keputusan pemberian beasiswa KIP-K di UIN SUSKA Riau menggunakan metode *Naïve Bayes* serta metodologi pengembangan *Agile* yang dilakukan oleh Al Ghani, dkk [6] menghasilkan tingkat akurasi sistem sebesar 88%. Di samping itu penelitian mengenai penggunaan Algoritma C4.5 untuk penentuan penerimaan beasiswa dengan menggunakan 125 data yang dilakukan oleh Suweleh, dkk [7] memiliki tingkat akurasi sebesar 92% dan dikembangkan menjadi aplikasi dengan metodologi *Waterfall* dan pemodelan DFD. Penelitian oleh Putra, dkk [8] meneliti mengenai penerapan algoritma *K-Nearest Neighbors* untuk membantu mendiagnosis penyakit stroke menghasilkan tingkat akurasi klasifikasi sebesar 88% serta metode pengembangan sistem *prototype* guna mengembangkan penelitian tersebut ke dalam sebuah aplikasi *mobile* dengan tingkat penerimaan pengguna sebesar 78%. Kemudian Yunus dan Pratiwi [9] meneliti penerapan algoritma *K-Nearest Neighbors* dalam sistem berbasis web menggunakan metode pengembangan *waterfall* untuk memprediksi status gizi balita di Puskesmas Cakranegara dengan hasil tingkat akurasi keberhasilan dalam menentukan status gizi balita mencapai 88,06%. Penelitian K-NN pun dilakukan oleh Qoiriyah, dkk [10] mengenai rancang bangun untuk menentukan jenis beasiswa dengan menggunakan algoritma *K-Nearest Neighbors* yang memiliki hasil akurasi 86% dan dikembangkan menjadi sistem pendukung keputusan dengan metodologi *Waterfall*. Penelitian dari Mulyati, dkk [11] membahas mengenai penerapan algoritma K-NN dalam memprediksi kelulusan ujian nasional di SMPN 2 Pagedangan berbasis desktop dengan data yang digunakan sebanyak 744 data. Sistem dikembangkan dengan metode SDLC, sistem ini memiliki tingkat akurasi 96,26% pada nilai  $k=7$  yang dapat memprediksi dan mengklasifikasikan dengan baik dan cepat. Penelitian rancang bangun sistem yang menerapkan model klasifikasi K-NN pun dilakukan oleh Said, dkk [12] dalam memprediksi kualitas air. Sistem prediksi tersebut dikembangkan berbasis web dengan metode pengembangan sistem SDLC. Penerapan model klasifikasi K-NN dalam sistem prediksi tersebut menghasilkan akurasi sebesar 85,24% dengan nilai  $k=3$  yang dapat memprediksi kualitas air dengan baik.

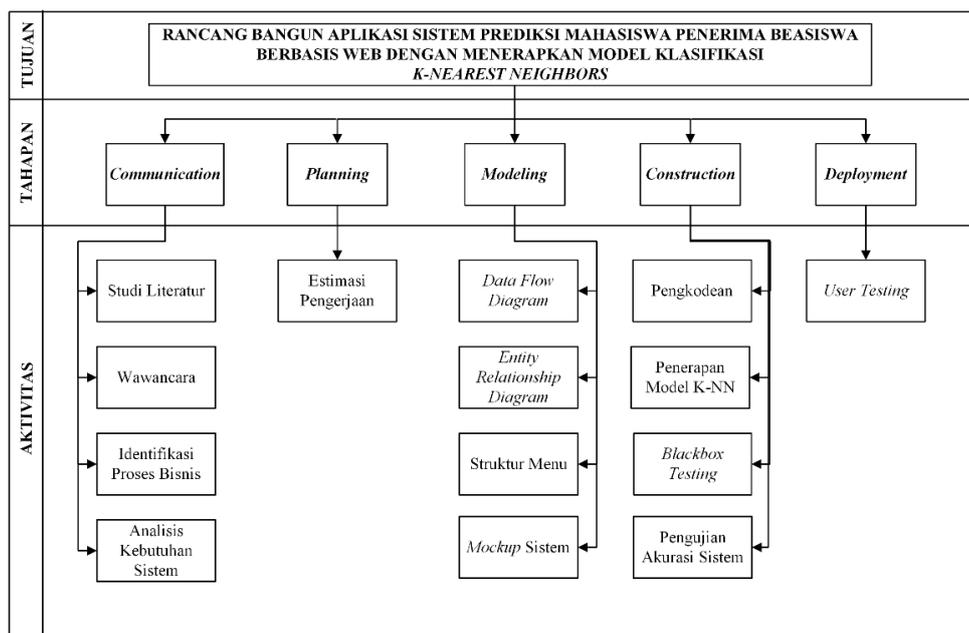
Berdasarkan rujukan penelitian yang telah diuraikan, penelitian yang dilakukan oleh Kurniadi, dkk [3] berhasil memodelkan klasifikasi calon penerima beasiswa dengan akurasi 90.12% dalam bentuk formula. Model tersebut menghasilkan akurasi 90.12% dengan data sebanyak 512 data, berdasarkan hasil penelitian Musu, dkk [13] nilai akurasi dipengaruhi oleh sebaran *dataset* sehingga dipilihlah model klasifikasi ini. Namun hasil penelitian tersebut belum dikembangkan menjadi sebuah aplikasi yang dapat digunakan. Oleh karena itu, tujuan penelitian ini berfokus untuk membangun sistem prediksi berbasis web dengan menerapkan model klasifikasi *K-Nearest Neighbors* yang diadaptasi dari penelitian sebelumnya oleh Kurniadi, dkk [3] yang dapat membantu menyeleksi calon mahasiswa penerima beasiswa dengan metode pengembangan sistem *Waterfall*, serta perbandingan tingkat akurasi yang dihasilkan sistem.

## 2. METODE PENELITIAN

Pada penelitian ini aplikasi sistem prediksi akan dibangun dengan menggunakan metodologi *waterfall* yang memiliki lima tahapan, yaitu *communication*, *planning*, *modelling*, *construction*, dan *deployment*. Model *waterfall* sendiri adalah model sekuensial linear yang menggunakan pendekatan sistematis dan sekuensial hadap pengembangan perangkat lunak [14] sehingga pengembangan sistem prediksi dapat terstruktur dan terukur dengan tahapan-tahapan yang jelas.

Model klasifikasi *K-Nearest Neighbors* (K-NN) akan diterapkan sebagai proses prediksi dimana nantinya sistem mampu mengklasifikasikan calon penerima beasiswa. *K-Nearest Neighbors* adalah metode klasifikasi terhadap sebuah objek berdasarkan data pembelajaran yang memiliki jarak paling dekat dengan objek tersebut. Data pembelajaran yang memiliki jarak terdekat dengan objek disebut sebagai tetangga terdekat. Jumlah dari data terdekat dinyatakan dengan *k*. Hasil klasifikasi K-NN diambil dari mayoritas kategori pada jumlah tetangga terdekat dari objek [15].

*Work Breakdown Structure* yang disajikan pada gambar 1 menunjukkan tujuan, tahapan, serta aktivitas yang akan dilakukan dalam membangun aplikasi sistem prediksi berbasis web untuk menyeleksi calon penerima beasiswa dengan menerapkan model klasifikasi K-NN.



Gambar 1. *Work Breakdown Structure*

Tahapan *communication* merupakan proses yang dilakukan dengan melakukan studi literatur, wawancara, identifikasi proses bisnis serta menganalisis kebutuhan sistem. Tahap *planning* dilakukan setelah aktivitas analisis kebutuhan sistem yaitu mengestimasi waktu pengerjaan sistem. Tahap *modeling* yaitu tahapan perancangan sistem dengan merancang *Entity Relationship Diagram*, *Data Flow Diagram*, struktur menu, dan *mockup* sistem. Tahap *construction* merupakan pengimplementasian perancangan ke dalam *coding*, penerapan model klasifikasi K-NN dalam sistem, *blackbox testing*, dan pengujian akurasi sistem. Tahap *deployment* yaitu penyerahan sistem kepada *user* untuk dilakukan uji sistem agar mendapatkan *feedback* secara langsung.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Communication

Tahapan *communication* merupakan proses inisiasi penelitian, terdapat beberapa aktivitas dalam tahapan *ini* di antaranya studi literatur, wawancara, identifikasi proses bisnis, dan analisis kebutuhan sistem.

##### 1) Studi literatur

Dari hasil studi literatur didapatkan bahwa penelitian dari Kurniadi, dkk [3] menghasilkan tingkat akurasi model klasifikasi sebesar 90,12% yang akan diterapkan sebagai sistem prediksi dalam pengembangan aplikasi penelitian ini. Selain itu, metode pengembangan terstruktur dirasa cocok digunakan sebagai metode pengembangannya karena aplikasi ini lebih berfokus kepada proses dan data sehingga dipilihlah metode *waterfall*.

Model klasifikasi K-NN ini menggunakan *dataset* pendaftar beasiswa KIP-K di Institut Teknologi Garut mulai dari tahun 2014-2019 sebanyak 512 data dengan 225 data berlabel layak dan 287 data berlabel tidak layak. Diketahui juga bahwa data yang digunakan memiliki banyak nilai *missing* sehingga penelitian sebelumnya *melakukan replace missing value* dengan menggunakan nilai rata-rata setiap atribut agar tidak ada data yang bernilai kosong. Selain itu, dilakukan proses penyeimbangan data agar tidak terdapat dominan kelas dengan menggunakan teknik SMOTE sehingga menghasilkan data sebanyak 574 data dengan 287 data layak dan 287 data tidak layak. Data yang digunakan memiliki karakteristik sebagai berikut:

Tabel 1. Karakteristik data

No	Atribut	Deskripsi
1	No. Pendaftaran (Integer)	Nomor yang dimiliki peserta berdasarkan pendaftaran di website KIPK Kemendikbud
2	Asal Sekolah (Polynomial)	Pendidikan terakhir dari peserta beasiswa
3	Pilihan Program Studi (Polynomial)	Jurusan yang dipilih oleh peserta beasiswa
4	Skor Nilai Seleksi (Real)	Nilai peserta hasil dari tes tulis
5	Skor Nilai Wawancara (Real)	Nilai peserta hasil dari tes wawancara
6	Skor Nilai Kondisi Ekonomi (Real)	Jumlah tanggungan dari jumlah pendapatan bulanan keluarga dibagi jumlah anggota keluarga
7	Skor Prestasi Akademik (Real)	Nilai gabungan dari peringkat kelas, rata-rata rapor dan nilai UAN
8	Skor Prestasi Non Akademik (Real)	Nilai gabungan dari raihan akademik baik lokal, regional, kabupaten, nasional
9	Skor Hasil <i>Survey</i> (Real)	Nilai gabungan dari penilaian tempat tinggal

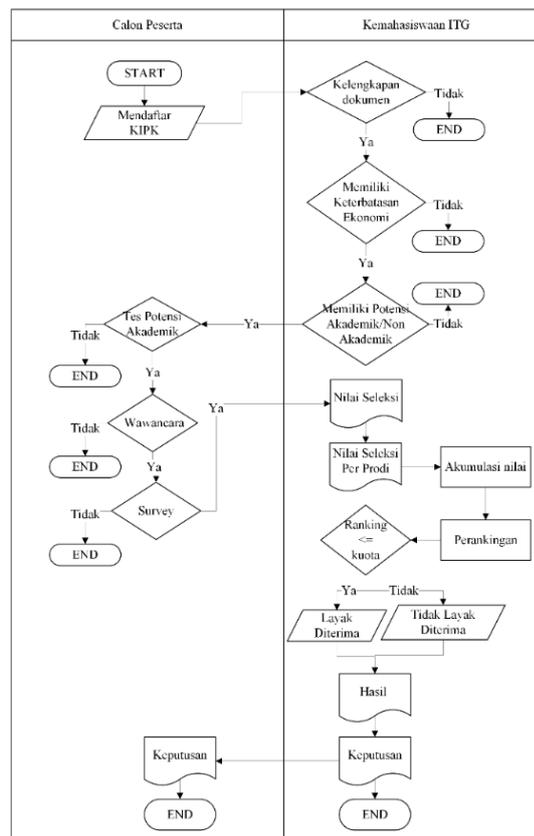
Akurasi model 90,12% tersebut dihasilkan dari pembagian data dengan rasio 70% data *training* dan 30% data *testing* dengan nilai  $k=3$  atau jumlah tetangga terdekat yang digunakan yaitu 3. Data memiliki 10 atribut *input* yaitu nomor pendaftaran, nama siswa, asal sekolah, pilihan program studi, skor nilai seleksi, skor nilai wawancara, skor nilai kondisi ekonomi, skor prestasi akademik, skor prestasi non akademik, skor hasil *survey* serta 1 atribut *output* yaitu label dengan 2 kategori layak dan tidak layak.

## 2) Wawancara

Berdasarkan hasil wawancara, bagian kemahasiswaan menyatakan bahwa permasalahan dari penerimaan KIP-K adalah jumlah kuota yang dimiliki tidak sebanding dengan jumlah pendaftar sehingga proses seleksi harus lebih ketat dan lebih tepat sasaran. Selain itu, tidak terdapat sistem khusus yang dapat membantu proses seleksi tersebut menjadikan proses seleksi memakan lebih banyak waktu dan tidak efisien. Oleh karena itu, perlu adanya sistem khusus yang dapat membantu mengklasifikasi dalam memprediksi kelayakan para calon penerima beasiswa yang akan diusulkan agar tepat sasaran dan proses penyeleksian dapat lebih cepat.

## 3) Identifikasi proses bisnis

Berikut merupakan proses bisnis yang sedang berjalan saat proses penyeleksian beasiswa KIP-K di ITG.



Gambar 2. Proses Bisnis yang sedang berjalan

Berdasarkan proses bisnis yang sedang berjalan, pada proses penyimpanan data dilakukan berulang sehingga tidak efisien, kemudian dikarenakan perhitungan nilai dilakukan secara manual dan tidak diseleksi secara otomatis sehingga jika terdapat banyak perndaftar maka data dari setiap pendaftar harus dilihat satu persatu menyebabkan penyeleksian membutuhkan waktu lama. Dari hal tersebut dibuatlah sistem prediksi yang dapat mengklasifikasi calon penerima beasiswa agar memudahkan bagian kemahasiswaan menyeleksi pendaftar dengan fokus utama terhadap pendaftar yang terklasifikasi layak.

#### 4) Analisis kebutuhan sistem

Aktivitas ini menjelaskan spesifikasi sistem yang dibutuhkan pada penelitian ini :

##### a. Spesifikasi Fungsional

- Sistem dapat diakses oleh Admin, Operator dan Peserta.
- Sistem dapat dilakukan *login logout*.
- Sistem dengan autentikasi Admin dapat mengelola data *training*, data *testing*, data pendaftar, data seleksi, data atribut program studi dan asal sekolah (*create, read, update, delete*) serta menguji tingkat akurasi sistem prediksi.
- Sistem dengan autentikasi Operator dapat memvalidasi data pendaftar, mengelola data seleksi, memprediksi kelayakan pendaftar KIP-Kuliah, serta menyaring pendaftar yang layak diterima.
- Sistem dengan autentikasi Peserta dapat melakukan registrasi untuk mendaftar KIP-Kuliah ITG serta mengakses sistem untuk informasi kelulusan dari pendaftaran KIP-Kuliah.

##### b. Spesifikasi Non Fungsional

###### a) Sisi Pengembang

Berikut beberapa syarat minimum dari sisi pengembang :

- Framework CodeIgniter* versi 4.3
- XAMPP versi 3.3.0 / PHP versi 8.1.12

- iii. Teks Editor *Visual Studio Code* versi 1.78.2
  - iv. Perangkat keras : Processor Intel(R) Celeron(R) N4000 1.10GHz, RAM 4 GB, dan Penyimpanan 500 GB.
- b) Sisi Pengguna
- Persyaratan minimum bagi pengguna yaitu hanya perlu memiliki sebuah *web browser*, baik itu Google Chrome, Firefox dan lainnya.

### 3.2 Planning

Tahapan ini hanya memiliki satu aktivitas yaitu mengestimasi waktu pengerjaan sistem yang akan dibangun. Aktivitas ini mengestimasi waktu dari setiap fitur, modul, atau menu dengan mempertimbangkan kompleksitas serta pengalaman tim pengembang. 1 hari = 8 jam kerja.

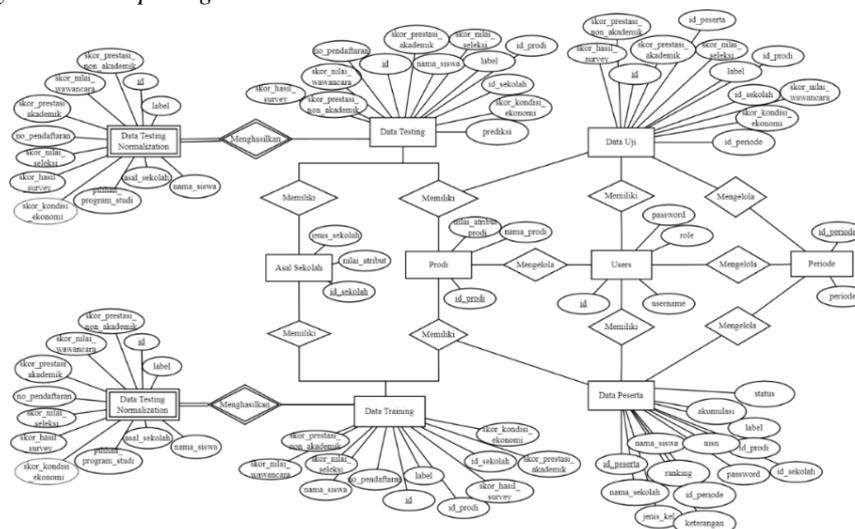
Tabel 2. Estimasi Pengerjaan

No.	Modul	Estimasi (Hari)
1	Login	1
2	Dashboard	2
3	Menu Data Prodi	3
4	Menu Data Asal Sekolah	3
5	Menu Data <i>Training</i>	5
6	Menu Data <i>Testing</i>	5
7	Normalisasi Data	6
8	<i>Euclidean</i>	4
9	<i>Import Data</i>	2
10	<i>Export Data</i>	1
11	<i>Sorting</i> Fitur	1
12	Menu Data Pendaftar	1
13	Menu Data Seleksi	2
14	Menu Hasil Prediksi	1
15	Menu Keputusan	1
16	Perankingan Data	1
17	Menu Periode	1
18	Menu Kelulusan	1
19	Menu Data Diri	1
<b>Total</b>		<b>42 Hari</b>

### 3.3 Modeling

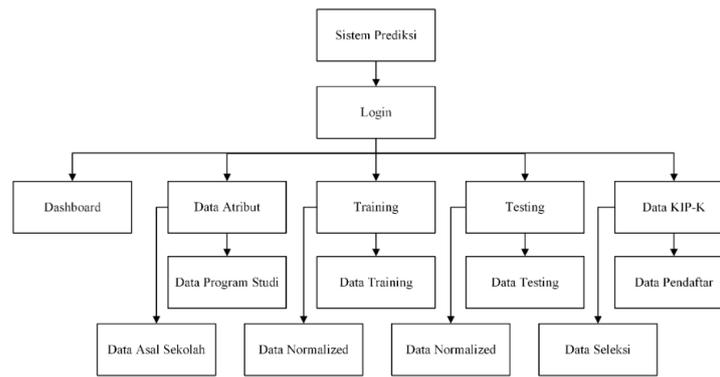
Tahapan ini memiliki beberapa aktivitas perancangan yaitu aktivitas membuat *Data Flow Diagram*, *Entity Relationship Diagram*, struktur menu dan *mockup* sistem.

#### 3.1 Entity Relationship Diagram



Gambar 3. ERD



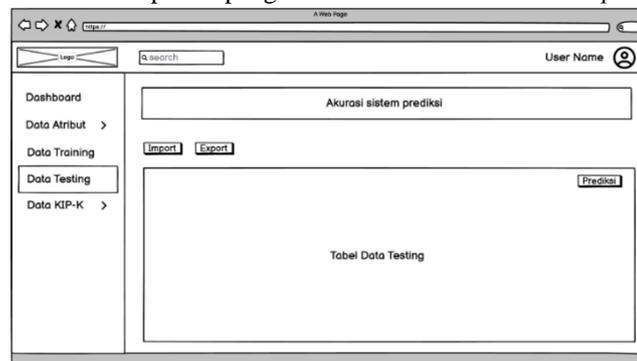


Gambar 6. Struktur Menu Admin

Gambar 6 memperlihatkan alur menu dari sistem prediksi bagi *user* dengan *role* admin. Menu admin berisikan data-data terkait dengan proses prediksi.

### 3.4 Mockup Sistem

*Mockup* sistem merupakan perancangan dari antarmuka sistem yang direncanakan akan dibuat dan diimplementasikan ke dalam proses pengkodean. Berikut contoh *mockup* sistem prediksi yang dibuat.



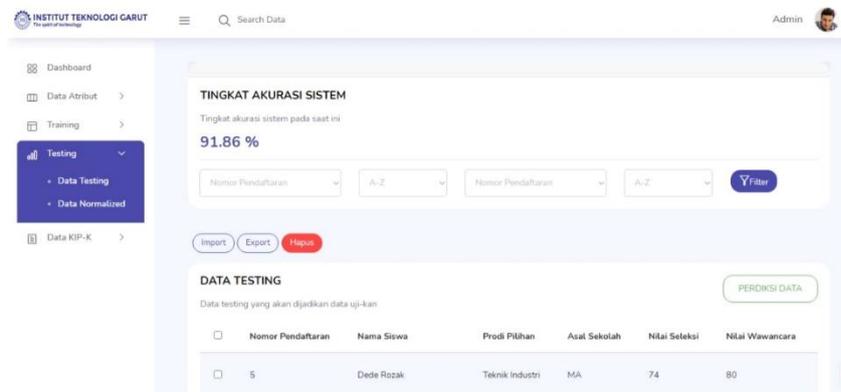
Gambar 7. Contoh Mockup Sistem

### 3.5 Construction

Pada tahapan ini, semua rancangan dan model klasifikasi K-NN akan diimplementasikan ke dalam bahasa pemrograman PHP *framework* CodeIgniter dengan struktur basis data MySQL dan terdapat dua pengujian sistem, yaitu *black box* testing untuk menguji fungsionalitas sistem dan pengujian akurasi sistem untuk menguji tingkat akurasi yang dimiliki oleh sistem.

#### 1) Pengkodean

Seperti yang telah disebutkan sebelumnya, aplikasi ini dibangun dengan menggunakan bahasa pemrograman PHP yang menggunakan *framework* CodeIgniter 4. Berikut merupakan hasil implementasi perancangan ke dalam proses pengkodean.



Gambar 8. Contoh Interface Sistem Prediksi

## 2) Penerapan Model K-NN

Penerapan model klasifikasi *K-Nearest Neighbors* dalam aplikasi ini digunakan untuk mengklasifikasikan kelayakan calon penerima beasiswa KIP-K. Untuk proses klasifikasi tersebut diperlukan beberapa rumus dan langkah yang harus diimplementasikan ke dalam bahasa pemrograman.

## a. Normalisasi data

Normalisasi dilakukan untuk menskalakan nilai dari suatu *variabel* menjadi lebih kecil antara -1 atau 1. Berikut rumus menormalisasikan data [16]. Normalisasi dilakukan pada seluruh data, baik data *training* maupun data *testing*.

$$v'_i = \frac{v_i - \min_A}{\max_A - \min_A} (\text{new\_max}_A - \text{new\_min}_A) + \text{new\_min}_A \quad (1)$$

Keterangan :

$v'_i$  : Nilai data baru hasil normalisasi

$v_i$  : Nilai data yang akan dinormalisasi

$\min_A$  : Nilai minimal data

$\max_A$  : Nilai maksimal data

$\text{new\_min}_A$  : Nilai minimal yang diharapkan dari proses normalisasi

$\text{new\_max}_A$  : Nilai maksimal yang diharapkan dari proses normalisasi

## b. Menghitung jarak antar data

Seperti yang telah dipaparkan sebelumnya pada metodologi penelitian, model klasifikasi K-NN mengklasifikasikan data berdasarkan tetangga terdekat dari data itu sendiri. Untuk menghitung kedekatan antar data dilakukan dengan menggunakan rumus *Euclidean distance* [15].

$$d_i = \sqrt{\sum_{i=1}^n (p_{2i} - q_{1i})^2} \quad (2)$$

Keterangan :

$p_i$  : sampel data

$q_i$  : data uji

$i$  : variabel data

$d_i$  : jarak

$n$  : dimensi data

## c. Tentukan jumlah tetangga terdekatnya

Jumlah tetangga terdekat atau  $k$  pada model klasifikasi yang menghasilkan akurasi 90,12% berada pada  $k=3$ . Maka sistem harus mencari 3 data *training* terdekat dengan data uji berdasarkan hasil perhitungan jarak dengan *Euclidean* sebelumnya. Dan ambil dominan klasifikasi dari tetangganya.

3) *Black Box Testing*

Aktivitas ini dilakukan untuk menguji fungsionalitas dari sistem untuk mencari kesalahan atau *error* pada sistem dengan menguji satu persatu fungsi pada sistem. Berikut merupakan hasil dari proses pengujian sistem dengan *black box*.

Tabel 3. *Black Box Testing*

No.	Pengujian	Skenario	Hasil yang diharapkan	Validasi
1	<i>Login</i>	Masukan <i>username</i> dan <i>password</i> Masuk ke halaman dashboard dengan benar		Valid
2	Menu data program studi	Klik menu data program studi	Menampilkan data program studi	Valid
3	Tambah data prodi	Isi semua data dengan benar	Data berhasil ditambahkan	Valid
4	Edit data program studi	Ubah data yang sudah ada	Data berhasil diubah	Valid
5	Hapus data program studi	Klik <i>icon</i> sampah lalu oke	Terdapat <i>alert</i> data berhasil dihapus	Valid

No.	Pengujian	Skenario	Hasil yang diharapkan	Validasi
6	Hapus <i>multiple</i> data	Pilih beberapa data, klik hapus	Terdapat <i>alert</i> data berhasil dihapus	Valid
7	Impor data training	Upload data dengan file excel	Data pada file berhasil ditambahkan	Valid
8	Ekspor data testing	Klik export pada menu data testing	Seluruh data testing terunduh dengan format file excel	Valid
9	Prediksi data testing	Klik prediksi data	Seluruh data testing memiliki klasifikasi berdasarkan prediksinya	Valid
10	Sorting data seleksi	Masukan bobot nilai dan kuota	Data seleksi tersorting berdasarkan nilai akumulasi tertinggi	Valid

#### 4) Pengujian Akurasi Sistem

Berdasarkan hasil pengujian pada penelitian sebelumnya, pembagian data 70% data *training* dan 30% data *testing* terhadap 574 data menghasilkan tingkat akurasi model klasifikasi yang mencapai 90,12%, menandakan bahwa model ini memiliki tingkat keberhasilan yang tinggi dalam memprediksi kelayakan penerima beasiswa. Maka dalam proses pengujian sistem prediksi ini pun menerapkan hal yang sama yaitu menguji sistem dengan jumlah data sebanyak 574 data dengan 70% data *training* dan 30% data *testing* dengan jumlah  $k$  yang digunakan yaitu 3 ( $k=3$ ), berikut rumus perhitungan akurasinya [17] :

$$\text{Jumlah data } \textit{testing} = 172 \text{ data}$$

$$\text{Jumlah prediksi benar} = 126 \text{ data}$$

$$\text{Akurasi} = \frac{\text{jumlah prediksi benar}}{\text{jumlah data uji}} \times 100\%$$

(3)

Hitung akurasi berdasarkan persamaan (3) :

$$\text{Akurasi} = \frac{126}{172} \times 100\% = 73,26\%$$

Dilihat dari hasil akurasi sistem, terdapat penurunan akurasi dengan model klasifikasinya. Setelah dilakukan analisis lebih lanjut, teknik *preprocessing* data pada penelitian sebelumnya tidak diketahui sehingga data yang digunakan harus divalidasi pada bagian kemahasiswaan. Berdasarkan pernyataan bagian kemahasiswaan, data yang memiliki *missing value* pada beberapa kategori tidak dapat digantikan hanya dengan nilai rata-rata tetapi data harus digantikan dengan nilai 0, *grade* tertentu, nilai minimal atau nilai maksimal. Kemudian data yang telah divalidasi diujikan kembali dengan langkah dan jumlah *dataset* yang sama seperti sebelumnya.

$$\text{Jumlah prediksi benar} = 158 \text{ data}$$

$$\text{Akurasi} = \frac{158}{172} \times 100\% = 91,86\%$$

Berdasarkan hasil tersebut tingkat akurasi setelah dilakukan validasi data meningkat menjadi 91,86%. Tingkat akurasi ini justru lebih baik jika dibandingkan dengan akurasi modelnya. Dari keseluruhan perbedaan tingkat akurasi ini menunjukkan bahwa tingkat akurasi bergantung kepada kualitas data yang digunakan.

### 3.6 Deployment

Tahapan *deployment* ini dilakukan dengan menyerahkan sistem untuk dilakukan pengujian agar peneliti dapat mendapatkan *feedback* mengenai kekurangan dari sistem. Aktivitas ini disebut sebagai *user testing*, pengujian dilakukan oleh bagian kemahasiswaan, berikut *feedback* yang diberikan :

- 1) Menyediakan fitur pembobotan nilai atribut untuk akumulasi data pendaftar agar dapat dilakukan proses perangkingan.

- 2) Menambahkan fitur kuota per program studi untuk menyaring data pendaftar yang memiliki akumulasi tertinggi berdasarkan program studi dengan maksimal data yang ditampilkan berdasarkan jumlah kuotanya.
- 3) Dapat memasukkan serta mengubah ranking pendaftar beasiswa.

### 3.7 Pembahasan

Hasil dari penelitian ini menunjukkan bahwa model klasifikasi *K-Nearest Neighbors* berhasil diterapkan ke dalam sistem prediksi berbasis web yang dibangun dengan menggunakan bahasa pemrograman PHP *framework* CodeIgniter 4 serta metode pengembangan sistem *waterfall* dapat menghasilkan akurasi yang tinggi. Dengan terdapat tiga user dalam sistem yaitu admin, operator dan peserta, selain dapat memprediksi kelayakan calon penerima beasiswa dengan menggunakan model klasifikasi, sistem ini dapat selalu mengecek tingkat akurasi dari sistem sehingga dapat diketahui apabila terjadi penurunan tingkat akurasi dari proses klasifikasinya, fitur pembobotan nilai atribut untuk menghitung akumulasi nilai agar dapat dilakukan proses perankingan calon penerima beasiswa, serta pengurutan data calon penerima beasiswa berdasarkan kelayakan dan akumulasi nilai tertinggi yang ditampilkan sesuai jumlah kuota per program studinya.

Sistem ini menghasilkan tingkat akurasi 91,86% dengan menggunakan  $k=3$ , akurasi ini lebih tinggi dibandingkan model klasifikasinya yaitu 90,12% dengan jumlah  $k$  yang sama, hal ini disebabkan perbedaan pada tahap *preprocessing data*. Meskipun begitu, hal ini selaras dengan hasil penelitian sebelumnya oleh Kurniadi, dkk [3] yang menyatakan bahwa dengan menerapkan model klasifikasi K-NN ini sistem dapat memiliki tingkat kurasi yang tinggi dalam memprediksi kelayakan calon penerima beasiswa. Selain itu, terjadinya perbedaan tingkat akurasi sistem dengan modelnya yang disebabkan oleh perbedaan *processing data* menunjukkan bahwa proses persiapan data berpengaruh pada tingkat akurasi yang dihasilkan, hal ini selaras dengan hasil penelitian dari Musu, dkk [13] yang menyatakan bahwa proses persiapan data harus diperhatikan agar mendapatkan *dataset* berkualitas yang mampu memberikan akurasi dengan nilai terbaik.

Dilihat dari segi urgensitas penelitian yang telah dipaparkan pada pendahuluan dan digambarkan pada proses bisnis, penelitian ini telah membuktikan bahwa dengan adanya sistem yang memiliki tingkat akurasi sebesar 91,86% ini dapat digunakan untuk mengatasi masalah bagian kemahasiswaan dalam menyeleksi kelayakan calon penerima beasiswa agar lebih cepat dengan bantuan fitur-fitur dalam sistem. Dimana sistem dapat berimplikasi pada proses adminitrasi dengan melakukan *copy* secara berulang untuk menyeleksi per prodi sehingga lebih efisien dan proses seleksi calon penerima bisa lebih berfokus pada peserta yang terprediksi kategori layak.

## 4. KESIMPULAN

Kesimpulan dari penelitian ini yaitu adanya aplikasi sistem prediksi berbasis web yang menerapkan model klasifikasi *K-Nearest Neighbors* yang diadaptasi dari hasil penelitian sebelumnya dengan tingkat akurasi prediksi tinggi, sehingga berimplikasi pada proses penyeleksian yang dapat mengurangi kerumitan administrasi karena *copy-paste* data berulang, serta lebih berfokus pada data terprediksi layak sehingga proses seleksi menjadi lebih selektif dan cepat. Kontribusi dari penelitian ini adalah terdapat alat yang dapat untuk membantu menyeleksi kelayakan calon penerima beasiswa KIP-K di Institut Teknologi Garut agar lebih tepat sasaran, serta menjadi referensi bagi sistem yang menerapkan model tertentu. Hasil pengujian akurasi penelitian ini menunjukkan bahwa kualitas *dataset* berpengaruh terhadap tingkat akurasi yang dihasilkan. Adapun saran untuk penelitian ini yaitu menambah cakupan fitur sistem untuk memantau mahasiswa KIP-K yang telah diterima di Institut Teknologi Garut.

## REFERENSI

- [1] I. Ilham, I. G. Suwijana, and N. Nurdin, "Sistem Pendukung Keputusan Penerimaan Beasiswa Pada SMK 2 Sojol Menggunakan Metode AHP," *Jurnal Elektronik Sistem Informasi dan Komputer*, vol. 4, no. 2, pp. 48-58, 2021.

- [2] Kemendikbud, “Pedoman Pendaftaran KIPK,” 2022. Accessed: Mar. 08, 2023. [Online]. Available: <https://kip-kuliah.kemdikbud.go.id/>
- [3] D. Kurniadi, F. Nuraeni, N. Abania, L. Fitriani, A. Mulyani, and Y. H. Agustin, “Scholarship Recipients Prediction Model using k-Nearest Neighbor Algorithm and Synthetic Minority Over-sampling Technique,” in *2022 12th International Conference on System Engineering and Technology (ICSET)*, 2022, pp. 89–94. doi: 10.1109/ICSET57543.2022.10010947.
- [4] A. Mulyani, D. Kurniadi, M. R. Nashrulloh, I. T. Julianto, and M. Regita, “The Prediction Of PPA And KIP-Kuliah Scholarship Recipients Using Naive Bayes Algorithm,” *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 4, pp. 821–827, 2022.
- [5] D. Haryanto, C. Ramdani, W. S. Wahidah, A. G. Dinia, and S. Oktaviani, “Perancangan Sistem Pendukung Keputusan Penentuan Penerimaan Beasiswa Menggunakan Metode Artificial Neural Network (ANN) di Institut Teknologi Telkom Purwokerto,” in *Conference on Electrical Engineering, Telematics, Industrial technology, and Creative Media (CENTIVE)*, 2019, pp. 103–114.
- [6] R. Al Ghani, A. F. Winanda, and M. L. Hamzah, “Sistem Pendukung Keputusan Pemberian Beasiswa KIP-K (Studi Kasus: ‘UIN SUSKA RIAU’),” in *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis*, 2022, pp. 236–239.
- [7] A. S. Suweleh, D. Susilowati, and H. Hairani, “Aplikasi Penentuan Penerima Beasiswa Menggunakan Algoritma C4. 5,” *Jurnal Bumigora Information Technology (BITE)*, vol. 2, no. 1, pp. 12–21, 2020.
- [8] D. A. Putra, M. D. S. Sanapiah, A. I. Hanifah, and T. Afirianto, “SEED (Stroke Disease Early Detection Application)–Rancang Bangun Aplikasi Mobile Berbasis Android Untuk Mendiagnosis Gejala Dini Penyakit Stroke Menggunakan K-Nearest Neighbor (K-NN),” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 6, no. 3, 2019.
- [9] M. Yunus and N. K. A. Pratiwi, “Prediksi Status Gizi Balita Dengan Algoritma K-Nearest Neighbor (KNN) di Puskesmas Cakranegara,” *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 4, no. 4, pp. 221–231, Feb. 2023, doi: 10.35746/jtim.v4i4.328.
- [10] L. Qoiriyah, H. L. Purwanto, and W. Setiyaningsih, “Rancang Bangun Sistem Pendukung Keputusan Penentuan Jenis Beasiswa Menggunakan Knn,” *RAINSTEK: Jurnal Terapan Sains & Teknologi*, vol. 1, no. 2, pp. 64–72, 2019.
- [11] S. Mulyati, S. M. Husein, and R. Ramdhan, “Rancang Bangun Aplikasi Data Mining Prediksi Kelulusan Ujian Nasional Menggunakan Algoritma (KNN) K-Nearest Neighbor Dengan Metode Euclidean Distance Pada SMPN 2 PAGEDANGAN,” *JIKA (Jurnal Informatika)*, vol. 4, no. 1, p. 65, Jan. 2020, doi: 10.31000/jika.v4i1.2288.
- [12] H. Said, N. H. Matondang, and H. N. Irmanda, “Perancangan Sistem Prediksi Kualitas Air Yang Dapat Dikonsumsi Dengan Menerapkan Algoritma K-Nearest Neighbor,” in *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*, 2022, pp. 158–168.
- [13] W. Musu, A. Ibrahim, and H. Heriadi, “Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4. 5,” in *SISITI: Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, 2021, pp. 186–195.
- [14] R. S. Pressman and B. R. Maxim, *Software Engineering : A Practitioner’s Approach*, Ninth Edition. New York: McGraw-Hill Education, 2020.
- [15] U. Hidayah and M. K. Agus Sifaunajah, *Cara Mudah Memahami Algoritma K-Nearest Neighbor Studi Kasus Visual Basic 6.0*. Lembaga Penelitian dan Pengabdian kepada Masyarakat Universitas KH. A. Wahab Hasbullah, 2019. [Online]. Available: <https://books.google.co.id/books?id=B6t-EAAAQBAJ>
- [16] U. Rofiqoh, R. Perdana, and M. Fauzi, “Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter dengan Metode Support Vector Machine dan Lexicon Based Features,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK)*, vol. 1, pp. 1725–1732, Aug. 2017.
- [17] D. Dahri, F. Agus, and D. M. Khairina, “Metode Naive Bayes Untuk Penentuan Penerima Beasiswa Bidikmisi Universitas Mulawarman,” *Jurnal Informatika Mulawarman*, vol. 11, no. 2, p. 29, 2016.