

SKRIPSI

**SISTEM REKOMENDASI PROGRAM STUDI UNIVERSITAS
KATOLIK PARAHYANGAN MENGGUNAKAN ALGORITMA
*COLLABORATIVE FILTERING***



Anugrah Jaya Sakti

NPM: 2016730053

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2020**

UNDERGRADUATE THESIS

**PARAHYANGAN CATHOLIC UNIVERSITY STUDY
PROGRAM RECOMMENDATION SYSTEM USING
COLLABORATIVE FILTERING ALGORITHM**



Anugrah Jaya Sakti

NPM: 2016730053

**DEPARTMENT OF INFORMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2020**

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

SISTEM REKOMENDASI PROGRAM STUDI UNIVERSITAS KATOLIK PARAHYANGAN MENGGUNAKAN ALGORITMA *COLLABORATIVE FILTERING*

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
Tanggal 15 Juni 2020



Anugrah Jaya Sakti
NPM: 2016730053

LEMBAR PENGESAHAN

SISTEM REKOMENDASI PROGRAM STUDI UNIVERSITAS KATOLIK PARAHYANGAN MENGGUNAKAN ALGORITMA *COLLABORATIVE FILTERING*

Anugrah Jaya Sakti

NPM: 2016730053

Bandung, 15 Juni 2020

Menyetujui,

Pembimbing

Husnul Hakim, M.T.

Ketua Tim Penguji

Anggota Tim Penguji

Rosa De Lima, M.T.

Natalia, M.Si.

Mengetahui,

Ketua Program Studi

Mariskha Tri Adithia, P.D.Eng

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

**SISTEM REKOMENDASI PROGRAM STUDI UNIVERSITAS KATOLIK
PARAHYANGAN MENGGUNAKAN ALGORITMA *COLLABORATIVE
FILTERING***

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
Tanggal 15 Juni 2020



Anugrah Jaya Sakti
NPM: 2016730053

ABSTRAK

Salah satu tahapan pendidikan setelah lulus SMA adalah melanjutkan studi ke perguruan tinggi. Salah satu hal yang perlu diperhatikan adalah memilih program studi apa yang akan dipilih. Berdasarkan [1], kesalahan dalam pemilihan program studi memiliki dampak negatif bagi mahasiswa, yaitu : masalah psikologi, akademis, dan relasional. Berdasarkan [2], pada tahun 2017 terdapat 1.437.425 mahasiswa baru, 6.924.511 mahasiswa terdaftar, dan 1.046.141 mahasiswa lulus. Salah satu cara yang dapat dilakukan untuk mengurangi kesalahan dalam pemilihan program studi adalah dengan menggunakan sistem yang dapat memberikan rekomendasi program studi yang tepat kepada calon mahasiswa.

Pada skripsi ini akan membahas sistem rekomendasi menggunakan algoritma *user-based collaborative filtering* untuk memberikan rekomendasi program studi berdasarkan *rating* tanpa memerlukan informasi mengenai item ataupun pengguna dan *K-Means* untuk pembentukan *cluster* untuk meningkatkan hasil akurasi. Pengujian dilakukan dengan *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE) untuk mengetahui akurasi sistem. Sistem yang dibangun menghasilkan rekomendasi lebih baik dengan *K-Means* dengan $k = 40$ untuk jurusan IPA dan $k = 30$ untuk jurusan IPS.

Kata-kata kunci: sistem rekomendasi, *user-based collaborative filtering*, *K-Means*, program studi Universitas Katolik Parahyangan

ABSTRACT

One of the stages of education after graduating from high school is to continue their studies to college. One important thing that need to be considered is choosing what majors to choose. Based on [1], wrong choice of majors will have a negative impact on the college students, whic is; psychological, academic, and relational problems. Based on [2], in 2017 there were 1,437,425 new students, 6,924,511 registered students, and 1,046,141 students graduated. One way that can be done to reduce this problem is to use a system that can provide appropriate majors recommendations to prospective students.

This thesis will discuss the recommendation system using the user-based collaborative filtering algorithm to provide majors recommendations based on rating without requiring information about items or users and K-Means for the formation of cluster to improve accuracy. Testing is done with Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to determine the accuracy of the system. The system built produces better recommendations with K-Means with k 40 for natural Sciences and k 30 for social Sciences.

Keywords: recommendation system, user-based collaborative filtering, K-Means, Parahyangan Catholic University study program

Skripsi ini saya persembahkan kepada ayah dan (Almh) ibu, terutama untuk (Almh) ibu. Semoga dengan skripsi ini bisa menjadi hadiah atas keinginan yang belum tersampaikan.

KATA PENGANTAR

Segala puji bagi Allah SWT yang telah memberikan rahmat dan karunia-Nya kepada penulis, sehingga penulis dapat menyelesaikan skripsi dengan judul "Rekomendasi Program Studi di Perguruan Tinggi untuk Siswa SMA". Shalawat dan salam senantiasa tercurah kepada Rasulullah SAW. Penyusunan skripsi ini dimaksudkan untuk memenuhi syarat guna mencapai gelar sarjana di Program studi Teknik Informatika, Fakultas Teknologi Informasi dan Sains, Universitas Parahyangan Bandung.

Penulis menyadari bahwa penulisan ini tidak dapat terselesaikan tanpa dukungan dari berbagai pihak, baik moril maupun materil. Oleh karena itu, penulis ingin mengucapkan ucapan terima kasih kepada :

1. Kedua orang tua, Bapak Ibrahim Warga Purawinata, (Almh) Ibu Dede Komariah, dan Syarif Jordan yang telah memberikan dukungan baik moril maupun materil serta doa.
2. Bapak Husnul Hakim, M.T. selaku dosen Pembimbing yang memberikan dukungan, kritik, saran, dan nasihat yang membangun dalam pelaksanaan penelitian ini maupun terhadap penulisan.
3. Ibu Rosa De Lima, M.T. dan Ibu Natalia, M.Si. selaku dosen penguji yang telah memberikan kritik dan saran yang membangun sehingga penelitian ini dapat selesai dengan baik.
4. Ibu Flaviana, M.T. selaku Wakil Dekan Bidang Kemahasiswaan dan Alumni yang telah membantu dalam pembuatan surat permohonan kepada Biro Admiristari Akademik (BAA) untuk kelancaran penelitian ini.
5. Seluruh Bapak/Ibu dosen Teknik Informatika yang telah memberikan pengetahuan dan pelajaran selama penulis berkuliah di Unpar.
6. Rizky Septika Wulandari yang selalu memberikan support.
7. Teman-teman seperjuangan selama perkuliahan, yaitu : Samuel, Dipo, dan Hashrul.
8. Serta seluruh teman-teman yang telah memberikan semangat dan dukungan yang tidak bisa disebutkan satu persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna dan menyakini bahwa kesempatan hanya milik Allah SWT, oleh karena itu apabila terdapat kesalahan dalam penelitian dan penulisan, penulis mengharapkan kritik dan saran yang dapat membantu penyempurnaan penelitian ini akan penulis terima dengan senang hati. Penulus berharap, semoga skripsi ini dapat bermanfaat bagi semua pihak yang membutuhkan.

Bandung, Juni 2020

Penulis

DAFTAR ISI

KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR	xix
DAFTAR TABEL	xxi
DAFTAR KODE PROGRAM	xxiii
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Batasan Masalah	3
1.5 Metodologi	3
1.6 Sistematika Pembahasan	4
2 LANDASAN TEORI	5
2.1 Sistem Rekomendasi	5
2.1.1 Fungsi Sistem Rekomendasi	5
2.1.2 Sumber Data dan Pengetahuan	6
2.1.3 Teknik Rekomendasi	7
2.1.4 <i>Collaborative Filtering</i>	8
2.1.5 Aplikasi dan Evaluasi	9
2.2 Analisis <i>Cluster</i>	11
2.2.1 Jenis-Jenis <i>Clustering</i>	11
2.2.2 Jenis-Jenis <i>Cluster</i>	12
2.2.3 K-Means	14
2.3 <i>Library</i> PHP-ML	15
2.3.1 Array Dataset	15
2.3.2 Random Split	15
2.4 Universitas Katolik Parahyangan	15
2.4.1 Program Studi	16
2.4.2 Syarat Masuk Program Studi	20
2.4.3 Karakteristik Program Studi	21
3 ANALISIS	23
3.1 <i>Preprocessing</i> Data Lulusan	23
3.2 Pemilihan Mata Pelajaran	25
3.3 Pemilihan Algoritma Sistem Rekomendasi	26
3.4 Contoh Perhitungan <i>Collaborative Filtering</i>	27
3.4.1 Contoh Perhitungan Kemiripan	28

3.4.2	Contoh Perhitungan Prediksi	31
3.5	Pemilihan Algoritma <i>Cluster</i>	32
3.6	Contoh Perhitungan <i>K-Means</i>	33
3.7	Contoh Perhitungan <i>Collaborative Filtering</i> dan <i>K-Means</i>	37
3.8	Contoh Perhitungan Metode Evaluasi Sistem Rekomendasi	38
3.9	Analisis Perangkat Lunak Sejenis	39
3.10	Analisis Kebutuhan Sistem	49
3.10.1	Diagram <i>Use Case</i>	49
3.10.2	Rancangan Basis Data	51
4	PERANCANGAN	53
4.1	Perancangan Fisik Basis Data	53
4.1.1	Perancangan Tabel	53
4.2	Perancangan Antarmuka	56
4.3	Perancangan <i>Class Diagram</i> dan Algoritma	58
4.3.1	Perancangan <i>Class Diagram</i>	59
4.3.2	Perancangan Algoritma	59
5	IMPLEMENTASI DAN PENGUJIAN	85
5.1	Lingkungan Implementasi	85
5.1.1	Lingkungan Perangkat Keras	85
5.2	Lingkungan Perangkat Lunak	85
5.3	Implementasi Tabel Basis Data	85
5.4	Implementasi Antarmuka	87
5.5	Pengujian Fungsional	90
5.5.1	Pengujian Fungsional Pemilihan Jurusan SMA	90
5.5.2	Pengujian Fungsional Pengisian Nilai	90
5.6	Pengujian Eksperimental	91
5.6.1	Sistem Rekomendasi tanpa <i>K-Means</i>	91
5.6.2	Sistem Rekomendasi dengan <i>K-Means</i>	93
5.6.3	Kesimpulan Hasil Pengujian	98
6	KESIMPULAN DAN SARAN	101
6.1	Kesimpulan	101
6.2	Saran	101
	DAFTAR REFERENSI	103
	A KODE PROGRAM	105

DAFTAR GAMBAR

2.1	Perbedaan cara <i>clustering</i>	11
2.2	Jenis-Jenis <i>Cluster</i>	13
2.3	Penggunaan K-Means untuk Menemukan Tiga <i>Cluster</i>	14
3.1	Data Mentah Lulusan Universitas Katolik Parahyangan	23
3.2	7 Dimensi Profil Siswa	40
3.3	Tampilan setelah registrasi atau <i>login</i>	41
3.4	Modul Pengenalan	43
3.5	Hasil Tes Kepribadian	44
3.6	Hasil Tes Minat	45
3.7	Hasil Tes Gaya Belajar	45
3.8	Hasil Tes <i>Personal Values</i>	46
3.9	Modul Potensi Diri	46
3.10	Modul Ukur Kemampuan Diri	48
3.11	Hasil Rekomendasi	48
3.12	Diagram <i>Use Case</i> Sistem Rekomendasi	51
3.13	Diagram ERD Sistem Rekomendasi	52
4.1	Halaman Index Sistem	56
4.2	Halaman Pengisian Nilai IPA	57
4.3	Halaman Pengisian Nilai IPS	57
4.4	Halaman Hasil Rekomendasi	58
4.5	Halaman Pengujian	58
4.6	<i>Class Diagram</i> Sistem Rekomendasi	59
5.1	Halaman Index Sistem	87
5.2	Halaman Index Pengisian Nilai IPA	88
5.3	Halaman Index Pengisian Nilai IPS	88
5.4	Halaman Hasil Rekomendasi IPA	89
5.5	Halaman Hasil Rekomendasi IPS	89
5.6	Halaman Pengujian	90
5.7	Visualisasi Pengujian Jurusan IPA Tanpa <i>K-Means</i>	92
5.8	Visualisasi Pengujian Jurusan IPS Tanpa <i>K-Means</i>	93
5.9	Visualisasi Pengujian Jurusan IPA	98
5.10	Visualisasi Pengujian Jurusan IPS	99

DAFTAR TABEL

2.1	Tabel syarat program studi	20
2.2	Tabel kriteria	21
3.1	Data Lulusan	24
3.2	Nilai Lulusan	24
3.3	Mata Pelajaran PMDK Program Studi	25
3.4	Data Siswa Sebelum di Konversi	27
3.5	Data Lulusan Sebelum di Konversi	27
3.6	Data Siswa Setelah di Konversi	28
3.7	Data Lulusan Setelah di Konversi	28
3.8	Perhitungan Kovarians Siswa dan Lulusan	29
3.9	Kuadrat Perbedaan Siswa	30
3.10	Kuadrat Perbedaan Lulusan	30
3.11	Kovarians dan Standar Deviasi Siswa dan Lulusan	31
3.12	Kemiripan Siswa dan Lulusan	31
3.13	Perhitungan Prediksi	32
3.14	Data Centroid 1	33
3.15	Data Centroid 2	33
3.16	Contoh Data Lulusan	33
3.17	Perhitungan Jarak <i>Centroid</i> 1 dengan Data Lulusan	34
3.18	Perhitungan Jarak <i>Centroid</i> 2 dengan Data Lulusan	34
3.19	Menempatkan Data ke <i>Centroid</i> Terdekat	35
3.20	Data Baru <i>Centroid</i> 1	35
3.21	Data Baru <i>Centroid</i> 2	36
3.22	<i>Final</i> Data <i>Centroid</i> 1	36
3.23	<i>Final</i> Data <i>Centroid</i> 2	36
3.24	Data Lulusan untuk <i>Centroid</i> 1	36
3.25	Data Lulusan untuk <i>Centroid</i> 2	36
3.26	Data Siswa untuk <i>K-Means</i>	37
3.27	Jarak Data Siswa dengan <i>Centroid</i> 1 dan 2	38
3.28	Kemiripan Siswa dengan Anggota <i>Centroid</i> 1	38
3.29	Perhitungan Prediksi	38
3.30	Tabel Data MAE dan RMSE	39
3.31	Pendefinisian Aktor	49
3.32	Pendefinisian <i>Use Case</i>	49
3.33	Skenario Memilih Jurusan SMA	50
3.34	Skenario Mengisi Nilai Rapor	50
4.1	Perancangan Tabel jurusan_sma	53
4.2	Perancangan Tabel fakultas	53
4.3	Perancangan Tabel program_studi	54
4.4	Perancangan Tabel mahasiswa	54
4.5	Perancangan Tabel mata_pelajaran	55

4.6	Perancangan Tabel nilai	55
5.1	Tabel Pengujian Fungsional Pemilihan SMA	90
5.2	Tabel Pengujian Fungsional Pengisian Nilai	91
5.3	Hasil Pengujian Jurusan IPA tanpa K-Means	91
5.4	Hasil Pengujian Jurusan IPS tanpa K-Means	92
5.5	Hasil Pengujian <i>K-Means</i> Jurusan IPA dengan nilai k 10	93
5.6	Hasil Pengujian <i>K-Means</i> Jurusan IPA dengan nilai k 20	94
5.7	Hasil Pengujian KMeans Jurusan IPA dengan nilai k 30	94
5.8	Hasil Pengujian KMeans Jurusan IPA dengan nilai k 40	95
5.9	Hasil Pengujian <i>K-Means</i> Jurusan IPS dengan nilai k 10	95
5.10	Hasil Pengujian <i>K-Means</i> Jurusan IPS dengan nilai k 20	96
5.11	Hasil Pengujian <i>K-Means</i> Jurusan IPS dengan nilai k 30	96
5.12	Hasil Pengujian <i>K-Means</i> Jurusan IPS dengan nilai k 40	97
5.13	Hasil Pengujian Jurusan IPA dengan Metode K-Means	98
5.14	Hasil Pengujian Jurusan IPS dengan Metode K-Means	99

DAFTAR KODE PROGRAM

5.1	Implementasi Tabel Jurusan SMA	86
5.2	Implementasi Tabel Fakultas	86
5.3	Implementasi Tabel Program Studi	86
5.4	Implementasi Tabel Mahasiswa	86
5.5	Implementasi Tabel Mata Pelajaran	86
5.6	Implementasi Tabel Nilai	86
A.1	filter_data.py	105
A.2	MahasiswaController.php	106
A.3	SiswaController.php	106
A.4	KMeansController.php	107
A.5	UserBasedModelController.php	110
A.6	PearsonCorrelationController.php	111
A.7	PredictionController.php	112
A.8	FakultasController.php	113
A.9	ProgramStudiController.php	113
A.10	PengujianController.php	113
A.11	PearsonCorrelationPengujianController.php	115
A.12	AccuracyController.php	116
A.13	Fakultas.php	117
A.14	Jurusan_SMA.php	117
A.15	Mahasiswa.php	117
A.16	Mata_Pelajaran.php	118
A.17	Nilai.php	118
A.18	Program_Studi.php	118
A.19	header.blade.php	118
A.20	index.blade.php	119
A.21	ipa.blade.php	120
A.22	ips.blade.php	121
A.23	result.blade.php	122
A.24	pengujian.blade.php	123
A.25	2020_02_07_141228_create_jurusan_sma_table.php	124
A.26	2020_02_07_141432_create_mata_pelajaran_table.php	125
A.27	2020_02_07_142008_create_fakultas_table.php	125
A.28	2020_02_07_143407_create_program_studi_table.php	125
A.29	2020_02_07_145514_create_mahasiswa_table.php	126
A.30	2020_02_07_145515_create_nilai_table.php	126
A.31	DatabaseSeeder.php	127
A.32	Fakultas.php	127
A.33	Jurusan_SMA.php	127
A.34	Jurusan_SMA.php	128
A.35	Jurusan_SMA.php	128

BAB 1

PENDAHULUAN

Pada bab ini berisi mengenai latar belakang dari pembuatan skripsi, rumusan masalah, tujuan pembuatan skripsi, batasan masalah yang digunakan dalam pembangunan perangkat lunak, metodologi yang digunakan selama menyusun skripsi, dan sistematika pembahasan pada setiap bab yang ada.

1.1 Latar Belakang

Salah satu tahapan pendidikan setelah lulus dari bangku sekolah menengah atas atau SMA adalah melanjutkan studi ke perguruan tinggi baik perguruan tinggi negeri ataupun swasta. Salah satu hal yang perlu diperhatikan saat akan melanjutkan studi di perguruan tinggi adalah program studi apa yang akan dipilih. Program studi adalah kesatuan rencana belajar sebagai pedoman penyelenggaraan pendidikan akademik dan/atau profesional yang diselenggarakan atas dasar suatu kurikulum serta ditujukan agar mahasiswa dapat menguasai pengetahuan, keterampilan, dan sikap sesuai dengan sasaran kurikulum ¹.

Berdasarkan [1], Terdapat beberapa kendala dan permasalahan yang akan dihadapi oleh mahasiswa untuk mencapai hasil yang memuaskan, yaitu : tidak mampu mengikuti perkuliahan, tidak optimis mengikuti perkuliahan, tidak tertarik dengan mata kuliah program studi setelah memilih program studi, dan salah memilih program studi. Hal tersebut dapat mengakibatkan mahasiswa tidak dapat mencapai IPK tinggi dan dapat mengalami *Droup Out*. Terdapat tiga dampak salah memilih program studi di perguruan tinggi, yaitu :

1. Masalah Psikologi

Mempelajari sesuatu yang tidak sesuai minat, bakat, dan kemampuan, merupakan pekerjaan yang sangat tidak menyenangkan, apalagi kalau itu bukan kemauan atau pilihan anak, tetapi desakan dari orang tua, pelajaran akan sulit dimengerti karena sudah ada *blocking* emosi. Memilih jurusan kuliah sesuai dengan saran teman atau *trend*, padahal tidak sesuai dengan minat diri juga punya dampak psikologis.

2. Masalah akademis

Masalah akademis mengakibatkan prestasi yang tidak optimal, banyak mengulang mata kuliah yang berdampak bertambahnya waktu dan biaya, kesulitan memahami materi, kesulitan memecahkan persoalan, ketidakmampuan untuk mandiri dalam belajar, dan akhirnya adalah rendahnya nilai indeks prestasi.

3. Masalah relasional

Salah memilih jurusan kuliah membuat anak tidak nyaman dan tidak percaya diri. Ia merasa tidak mampu menguasai materi perkuliahan sehingga ketika hasilnya tidak memuaskan, merasa kurang percaya diri karena merasa dirinya kurang pandai, menjaga jarak dengan teman lain, semakin pendiam, menarik diri dari pergaulan, dan lebih senang mengurung diri.

¹Keputusan Menteri Pendidikan Nasional Nomor 232/U/2000 Pasal 1 Ayat 5

Berdasarkan [2], pada tahun 2017 terdapat 1.437.425 mahasiswa baru, 6.924.511 mahasiswa terdaftar, dan 1.046.141 mahasiswa lulus. Jumlah mahasiswa *Drop Out* pada tahun 2017 adalah 195.176 dengan persentase pada Perguruan Tinggi Negeri (PTN) sebesar 96% dan pada Perguruan Tinggi Swasta (PTS) sebesar 4%.

Untuk dapat mengurangi kesalahan dalam memilih program studi, ada banyak cara yang bisa dilakukan. Cara-cara tersebut antara lain adalah dengan membuat sebuah sistem yang dapat memberikan rekomendasi program studi yang tepat kepada calon mahasiswa. Sistem seperti ini dikenal dengan sistem rekomendasi. Berdasarkan [3], Sistem rekomendasi adalah alat dan teknik perangkat lunak yang menyediakan saran untuk item yang akan digunakan oleh pengguna. Saran terkait dengan berbagai proses pengambilan keputusan, seperti barang apa yang akan dibeli, musik apa yang akan didengarkan, atau berita online apa yang akan dibaca. Sistem rekomendasi berfokus pada item tertentu dan ditujukan untuk individu atau personal. Beberapa teknik yang biasa digunakan pada sistem rekomendasi, yaitu : *Content-based, Collaborative Filtering, Demographic, Knowledge-based, Community-based, Hybrid recommender systems*. Pada skripsi ini, teknik yang akan digunakan adalah *Collaborative Filtering*.

Collaborative Filtering [3] merupakan teknik yang merekomendasikan item yang sesuai dengan kebutuhan pengguna berdasarkan *rating* tanpa memerlukan informasi mengenai item ataupun pengguna, contoh informasi yang dimaksud adalah deskripsi mengenai item atau pengguna. Secara sederhana, *Collaborative Filtering* menghitung kemiripan antara pengguna aktif dengan beberapa pengguna lain yang memiliki selera atau minat yang serupa. Untuk menghitung kemiripan digunakan metode *Pearson Correlation Coefficient*. *Pearson Correlation Coefficient* bekerja dengan cara menghitung korelasi antara dua atribut dari masing-masing pengguna yang sedang dibandingkan. Atribut adalah sifat atau karakteristik dari tiap entitas maupun tiap relationship². Semakin tinggi nilai korelasi yang dihasilkan maka mengidentifikasi kedua pengguna memiliki kemiripan yang cukup tinggi.

Pada skripsi ini akan dibangun sebuah perangkat lunak sistem rekomendasi yang dapat memberikan rekomendasi item berupa program studi yang sesuai dengan minat siswa SMA. Terdapat dua teknik pada *Collaborative Filtering* [3], yaitu : metode *Neighborhood* dan *Latent Factor*. Metode *Neighborhood* fokus kepada relasi antara item atau pengguna, Terdapat dua pendekatan yaitu : *user-based* dan *item-based*. Metode *Latent Factor* merupakan faktorisasi matriks (SVD), terdiri dari pendekatan alternatif dengan mengubah item dan pengguna ke ruang faktor laten yang sama. Sistem rekomendasi ini akan menggunakan algoritma *Collaborative Filtering* dengan model *Neighborhood* dengan pendekatan *user-based*. *User-based* memprediksi berdasarkan kemiripan *rating* pengguna dengan item. *Rating* di dalam kasus ini adalah indeks prestasi kumulatif (IPK) lulusan pada program studi S1 yang berada di Universitas Katolik Parahyangan. Hasil penilaian capaian pembelajaran lulusan pada akhir program studi dinyatakan dengan indeks prestasi kumulatif (IPK)³ Selain menggunakan algoritma. *Collaborative Filtering* sistem ini akan menggunakan algoritma *K-Means* untuk pengelompokan data lulusan yang ada untuk meningkatkan hasil rekomendasi.

Pengujian akan dilakukan dengan membagi data lulusan secara acak menjadi *train set* sebesar 70% dan *test set* sebesar 30%, dimana data pada *train set* adalah data yang digunakan untuk membuat model *user based collaborative filtering* dan *test set* adalah data yang digunakan untuk menghitung akurasi dari hasil rekomendasi yang diberikan oleh sistem. Perhitungan akurasi akan digunakan berbagai ukuran, yaitu : *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan eksekusi waktu program (Time) dalam detik. Penggunaan MAE dan RMSE dilakukan dengan membandingkan IPK sebenarnya dengan IPK hasil prediksi dari sistem rekomendasi. IPK yang dibandingkan adalah IPK yang sesuai dengan program studi lulusan di dalam *test set*. Sebagai contoh, jika lulusan berasal dari Program Studi Akuntansi, maka hasil prediksi berupa IPK untuk Program Studi Akuntansi. Pengujian dilakukan untuk hasil rekomendasi tanpa *K-Means* dan

²Mulia Rahmayu, RANCANG BANGUN SISTEM INFORMASI NILAI UJIAN SISWA SMP NEGERI 3 BUMIAYU BERBASIS WEB, 2015, 3

³Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor Tahun 2014 tentang Standar Nasional Pendidikan Tinggi Pasal 23 ayat 5

dengan *K-Means*.

1.2 Rumusan Masalah

Berikut adalah rumusan masalah dari penulisan skripsi :

1. Bagaimana cara menilai kecocokan seorang calon mahasiswa terhadap suatu program studi ?
2. Bagaimana membangun perangkat lunak untuk memberikan rekomendasi program studi di perguruan tinggi yang cocok untuk calon mahasiswa ?
3. Bagaimana kualitas hasil rekomendasi dari perangkat lunak yang dibangun ?

1.3 Tujuan

Tujuan dari penulisan skripsi ini adalah sebagai berikut :

1. Mempelajari cara menilai kecocokan seorang mahasiswa terhadap suatu program studi.
2. Membangun perangkat lunak untuk memberikan rekomendasi program studi di perguruan tinggi yang cocok dengan calon mahasiswa.
3. Menguji hasil rekomendasi dari perangkat lunak yang sudah dibangun.

1.4 Batasan Masalah

Mengingat banyaknya perguruan tinggi dan program studi yang ada di Indonesia, maka perlu adanya batasan masalah yang jelas mengenai apa yang dibuat dan diselesaikan dalam penulisan skripsi ini. Berikut merupakan batasan-batasan masalah pada skripsi ini :

1. Program studi yang dijadikan rekomendasi adalah 15 program studi S1 di Universitas Katolik Parahyangan (UNPAR) Bandung.
2. Data yang akan digunakan adalah data lulusan S1 Universitas Katolik Parahyangan yang masuk melalui jalur Penelusuran Minat dan Kemampuan (PMDK) pada tahun 2013-2018.
3. Hanya menggunakan nilai mata pelajaran Matematika, Bahasa Indonesia, Bahasa Inggris, Pendidikan Kewarganegaraan, Fisika, dan Kimia sebagai atribut.
4. Sistem rekomendasi diperuntukan untuk siswa SMA kelas XI atau XII dengan jurusan IPA atau IPS yang ingin melanjutkan studi S1 di Universitas Katolik Parahyangan.
5. Sistem rekomendasi yang dibangun tidak memiliki fitur untuk admin.

1.5 Metodologi

1. Melakukan studi literatur mengenai sistem rekomendasi.
2. Mempelajari mengenai berbagai program studi dan karakteristiknya.
3. Mempelajari metode yang dapat digunakan untuk menghitung tingkat kecocokan calon mahasiswa dengan program studi.
4. Mempelajari *framework* yang dapat membantu pembangunan perangkat lunak. Dalam skripsi ini, akan digunakan Laravel dan Bootstrap. Karena itu, kedua *framework* ini akan dipelajari.

5. Menganalisis hal-hal yang mempengaruhi kecocokan program studi dengan calon mahasiswa.
6. Melakukan perancangan basis data, antarmuka, algoritma, dan *class diagram*.
7. Membangun perangkat lunak sesuai dengan analisis dan perancangan yang dilakukan.
8. Melakukan pengujian kualitas hasil rekomendasi perangkat lunak yang dibangun.
9. Menulis dokumen skripsi.

1.6 Sistematika Pembahasan

1. Bab 1 menjelaskan mengenai latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian, dan sistematika pembahasan untuk sistem rekomendasi program studi di perguruan tinggi untuk anak SMA.
2. Bab 2 menjelaskan mengenai sistem rekomendasi dengan menggunakan *Collaborative Filtering*, teknik perhitungan kemiripan menggunakan *Pearson Correlation Coefficient*, teknik *clustering* menggunakan *K-Means*, *Library PHP* yang digunakan, dan Program Studi yang berada di Universitas Katolik Parahyangan.
3. Bab 3 menjelaskan mengenai *preprocessing* data, mata pelajaran apa yang akan digunakan sebagai atribut, algoritma sistem rekomendasi yang akan digunakan, contoh perhitungan *collaborative filtering* dan evaluasi, algoritma *cluster* yang digunakan, contoh perhitungan *K-Means* analisis perangkat lunak terkait, dan kebutuhan sistem yang akan dibangun.
4. Bab 4 menjelaskan perancangan fisik basis data, antarmuka, algoritma, *class diagram*, dan algoritma yang akan digunakan sistem.
5. Bab 5 berisi implementasi dari analisis dan rancangan yang sudah dilakukan pada Bab 3 dan 4 untuk sistem rekomendasi program studi Universitas Katolik Parahyangan.
6. Bab 6 berisi kesimpulan yang dapat diambil oleh penulis dan saran untuk memberikan hasil rekomendasi yang lebih baik.