

Sistem Penunjang Keputusan Penentuan Dosen Pembimbing Menggunakan Algoritma Naive Bayes Studi Kasus STIKI Malang

Febri Yohanes Aldi Wicaksono¹, Anita², Chaulina Alfianti Oktavia³

^{1,2,3} Sekolah Tinggi Informatika dan Komputer Indonesia Malang

¹141111016@mhs.stiki.ac.id, ²ant@stiki.ac.id, ³chaulina@stiki.ac.id

ABSTRAK

Ketepatan dan kesesuaian dalam pemilihan dosen pembimbing dan co pembimbing tugas akhir bagi mahasiswa mempunyai peranan penting bagi keberhasilan tugas akhir mahasiswa. Mahasiswa sangat memerlukan dosen pembimbing yang dapat memberikan arahan dan masukan bagi mahasiswa yang sedang mengerjakan tugas akhir serta memahami judul dan tema studi kasus tugas akhir mahasiswa. Saat ini Kaprodi masih menggunakan sistem manual dalam menentukan dosen pembimbing tugas akhir. Hal tersebut memerlukan waktu yang lama dalam proses penentuan dosen pembimbing tugas akhir bagi mahasiswa yang bersangkutan, salah satu faktornya adalah kaprodi harus meneliti satu persatu judul sesuai calon dosen pembimbing yang mempunyai pengalaman didalam judul yang akan di angkat. Untuk itu diperlukan sebuah sistem yang dapat membantu Kepala Program Studi (KAPRODI) yang dapat merekomendasikan dosen pembimbing dan co pembimbing yang sesuai dengan kompetensi judul tugas akhir mahasiswa. Sistem rekomendasi dosen pembimbing yang dibangun memanfaatkan algoritma Naive Bayes *Classifier* dan *cosine similarity* sebagai penentu rekomendasi dosen pembimbing dan co pembimbing Naive Bayes *Classifier* merupakan teknik prediksi berbasis *probabilistic* sederhana yang berdasar pada penerapan teorema bayes atau aturan bayes dengan asumsi independensi (ketidaktergantungan) yang kuat sedangkan algoritma *cosine similarity* merupakan metode yang digunakan untuk menghitung *Similarity* (tingkat kesamaan) antar dua buah objek. Rekomendasi dosen pembimbing dan co pembimbing didasarkan pada kompetensi dosen pembimbing dan judul tugas akhir yang pernah dibimbing oleh dosen sebelumnya sehingga sistem dapat merekomendasi dosen pembimbing yang sesuai dan berkompoten. dengan acuan data training judul tugas akhir didapat hasil yang dapat membantu dalam proses penentuan dosen pembimbing bagi kaprodi STIKI Malang dan mahasiswa.

Kata Kunci : sistem penunjang keputusan, naïve bayes, dosen pembimbing, STIKI Malang

ABSTRACT

Accuracy and suitability in the selection of supervisors and co-supervisors for the final project have an important role for the success of the student's final project. Students really need a supervising lecturer who can provide direction and input for students who are working on their final project and understand the title and theme of their final assignment case studies. Currently the Head of Study Program is still using a manual system in determining the final project supervisor. This requires a long time in the process of determining the final project supervisor for the student concerned, one of the factors is the Head of Study Program must examine one by one the title according to the prospective supervisor who has experience in the title to be appointed. For this reason, a system that can help the Head of the Study Program (KAPRODI) is needed to recommend supervisors and co-supervisors who are in accordance with the competencies of the students' final assignment titles. The supervisor lecturer recommendation system that was built utilizes the Naive Bayes Classifier algorithm and cosine similarity as determinants of the Naive Bayes Classifier supervisor and co-algorithm is a simple probabilistic-based prediction technique based on the application of the Bayes theorem or Bayes rules with strong independence assumptions while the algorithm cosine similarity is a method used to calculate Similarity (level of similarity) between two objects. The recommendation of the supervisor and co-supervisor is based on the competence of the supervisor and the title of the final project that has been guided by the previous lecturer so that the system can recommend a suitable and competent supervisor. with reference to the training data of the final project title, the results obtained can help in the process of determining the supervisor for the STIKI Malang study program head and students.

Keywords : decision support system, naïve bayes, supervisor lecturer, STIKI Malang

1. PENDAHULUAN

STIKI Malang adalah salah satu sekolah tinggi yang berada di kota Malang berdiri pada tahun 1985. Saat ini STIKI Malang mempunyai 4 program studi yaitu Teknik Informatika (S1), Desain Komunikasi Visua (S1), Sistem Informasi

(S1), dan Manajemen Informasi (D3). Salah satu syarat wajib untuk kelulusan bagi mahasiswa STIKI adalah tugas akhir. Tugas akhir adalah karya tulis ilmiah mahasiswa yang merupakan hasil penelitian yang dilakukan untuk memberi solusi atas permasalahan yang diteliti atau berupa kajian ilmiah terhadap suatu konsep. Untuk

menyelesaikan tugas akhir mahasiswa akan di bimbing dan diarahkan oleh dosen pembimbing dan co pembimbing, penentuan dosen pembimbing dilakukan oleh kepala program studi (KAPRODI) dengan mengikuti aturan yang berlaku.

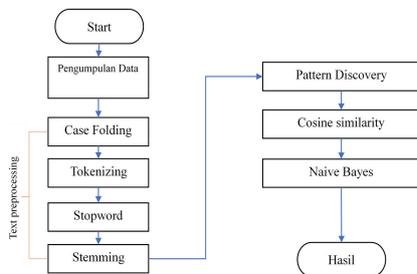
Hal tersebut memerlukan waktu yang lama dalam proses penentuan dosen pembimbing tugas akhir bagi mahasiswa yang bersangkutan, salah satu faktornya adalah kaprodi harus meneliti satu persatu judul sesuai calon dosen pembimbing yang mempunyai pengalaman didalam judul yang akan diangkat. Untuk mengatasi hal tersebut maka diperlukan sebuah sistem yang dapat membantu mengkategorikan dan mengklasifikasi dokumen berdasarkan pengalaman dari dosen pembimbing.

Oleh karena itu pada penelitian ini akan di bangun sebuah aplikasi menggunakan metode *Naive Bayes* dan *Cosine Similarity* yang dapat mengklasifikasi dan menampilkan hasil rekomendasi dosen pembimbing yang sesuai dengan judul tugas akhir dan diharapkan dapat membantu Kaprodi dalam menentukan dosen pembimbing tugas akhir.

2. ANALISA DAN PERANCANGAN

Penerapan Algoritma pada Sistem

Tahapan – tahapan algoritma yang akan dilakukan untuk memproses data training dan uji Sistem Penunjang Keputusan Penentuan Dosen Pembimbing menggunakan algoritma *Naive Bayes* adalah sebagai berikut



Gambar 1. Tahapan Penerapan Algoritma

Pengumpulan Data

Pengumpulan data dilakukan dengan mengumpulkan data abstraksi sebagai data training dan data judul penelitian ini sebagai data uji. Data tersebut dikumpulkan secara *manual* dengan cara mengetik ulang data abstraksi yang ada di perpustakaan STIKI Malang.

Tabel 1. Contoh *Sample Data*

No	Judul	Dosen Pembimbing	Co Pembimbing
1	Game 2D Sinau Song Untuk Belajar Lagu Daerah Dan Nasional dengan	Dr. Eva Handriyantini, S.Kom., M.MT	Chaulina Alfianti Oktavia, S.Kom., M.T

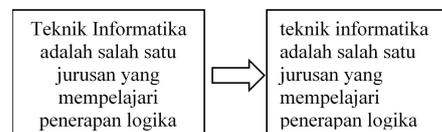
2	Construct 2 Sistem Informasi Pengelolaan Data Lembur Dan Cuti Pegawai Pada Kantor Pengelola Data Elektronik Kabupaten Malang Berbasis Web	Sugeng Widodo.,M.Kom	-
3	Sistem Pakar Pendekteksi Kerusakan Sepeda Motor Non-Injeksi Menggunakan Depth - First Search Dan Forward Chaining Berbasis Android	Diah Arifah P, S.Kom.,M.T	Chaulina Alfianti Oktavia, S.Kom., M.T
4	Sistem Informasi Daerah Rawan Bencana Alam Berbasis Web	Anita, S.Kom, M.T.	Johan Ericka W P, M.Kom
5	Sistem Informasi Pengelolaan Persewaan Alat Musik Dan Studio Musik AS	Koko Wahyu Prasetyo, S.Kom.,M.T.I	Yekti Asmoro Kanthi, S.Si., M.A.B

Tabel 2. Contoh Studi Kasus

Q	Judul	Dosen Pembimbing	Co. Pembimbing
Q	Rancang Bangun Sistem Keamanan Smartphone Berbasis Android	?	?
Target	Rancang Bangun Sistem Keamanan Smartphone Berbasis Android	Daniel Rudiaman S, ST., M.Kom	Chaulina Alfianti Oktavia, S.Kom., M.T

Case Folding

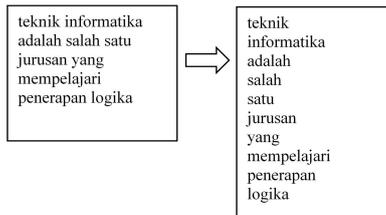
Tahap *Case Folding* digunakan mengubah huruf pada kalimat menjadi non kapital[1].



Gambar 2. *Case Folding*

Tokenizing

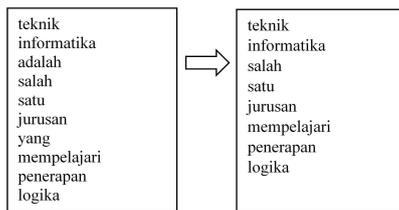
Tahap *tokenizing* adalah tahap pemotongan *string input* berdasarkan tiap kata yang menyusunnya.



Gambar 3. Tokenizing

Filtering/Stopword

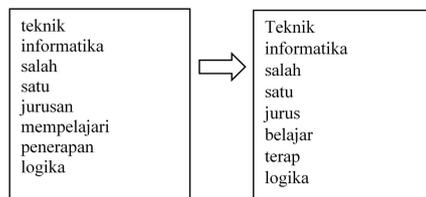
Tahap *filtering/stopword* adalah tahap pengambilan kata-kata penting dari hasil token.



Gambar 4. Filtering/Stopword

Stemming

Tahap *Stemming* adalah tahap untuk memperkecil jumlah indeks yang berbeda dari suatu dokumen serta melakukan pengelompokan kata lainya yang memiliki kata dasar dan arti yang serupa namun mempunyai bentuk yang berbeda karena mendapatkan imbuhan yang berbeda.



Gambar 5. Stemming

TF-IDF (Term Frequency-Inverse Document Frequency)

Dari sinilah hasil yang paling utama dalam *Text Mining* yaitu untuk menentukan bobot nilai yang telah diproses menggunakan TF/IDF. IDF dapat memperbaiki nilai *precision*, karena IDF mengkhususkan untuk fokus pada sebuah term dalam keseluruhan dokumen. IDF dapat memperbaiki nilai *precision*, karena IDF dapat mengkhususkan untuk fokus pada sebuah term dalam keseluruhan dokumen. kombinasi antara TF dan IDF untuk menghitung bobot term menunjukkan bahwa gabungan keduanya menghasilkan performansi yang lebih baik. Kombinasi bobot dari sebuah *term t* pada teks d didefinisikan sebagai berikut :

Pertama menentukan berapa TF yang terdapat pada dokumen dengan menggunakan rumus.

$$W(d, t) = T(d, t) F W(d, t) = T(d, t)$$

Term	sistem	game	Belajar	Data	lagu	informasi	artificial	terapan	algoritma	tunjang
D1	0	1	1	1	1	0	0	0	0	0
D2	1	0	0	2	0	1	0	0	0	0
D3	1	0	0	0	0	0	0	0	0	0
D4	1	0	0	0	0	1	0	0	0	0
D5	1	0	0	0	0	1	0	0	0	0

Gambar 6. Hasil TF

Nilai TF sudah diketahui maka akan di lanjutkan dengan mencari nilai IDF dalam dokumen tersebut dengan menggunakan.

$$IDF(t) = \log \left(\frac{N}{df(t)} \right)$$

Term	sistem	game	Belajar	Data	lagu	informasi	artificial	terapan
D1	0	1,113943352	1,11394	1,1139	1,113943352	0	0	0
D2	1,113943352	0	0	0,8451	0	1,113943352	0	0
D3	1,113943352	0	0	0	0	0	0	0
D4	1,113943352	0	0	0	0	1,113943352	0	0
D5	1,113943352	0	0	0	0	1,113943352	0	0

Gambar 7. Hasil IDF

Cosine Similarity

Langkah awal perhitungan adalah menghitung jarak antara dokumen (D) dengan dokumen contoh kasus (Q) dengan rumus sehingga didapatkan hasil :

$$CosSim(d_i, q_i) = \frac{q_i \cdot d_i}{|q_i| |d_i|} = \frac{\sum_{j=1}^t (q_{ij} \cdot d_{ij})}{\sqrt{\sum_{j=1}^t (q_{ij})^2 \cdot \sum_{j=1}^t (d_{ij})^2}}$$

Dokumen	CosD1	CosD2	CosD3	CosD4
SIMILARITY	0,02196	0,13239	0,248078	0,158716

Gambar 8. Perhitungan Cosine Similarity

Keterangan :

Cos D1 :

$$\sqrt{0 * 0 + 1 * 1 + 1 * 1 + 1 * 1 + 1 * 1 + 1 * 1 + \dots + 0 * 0} = 2,76487$$

Dimana D1 adalah rumus perhitungan vektor dokumen pada hasil IDF.

Naive Bayes

Setelah nilai dari *Cosine Similarity* antar dokumen ditemukan maka hasil dari tiap perhitungan akan dikategorikan sesuai dengan

nama dosen yang tertera sesuai judul dokumen kemudian akan dilakukan perhitungan nilai probabilitas pada tiap dosen pembimbing dan co pembimbing.

Dokumen	Nama dosen Pembimbing	Probabilitas Dosen pembimbing (Pdosen)
Pdosen-D1	Dr. Eva Handriyantini, S.Kom.,M.MT	0,010862636
Pdosen-D2	Sugeng Widodo.,M.Kom	0,116914394
Pdosen-D3	Diah Arifah P, S.Kom.,M.T	0,198768076
Pdosen-D4	Anita, S.Kom, M.T.	0,136975989
Pdosen-D5	Koko Wahyu Prasetyo, S.Kom.,M.T.I	0,25062916

Gambar 9. Perhitungan Cosine Similarity

Keterangan Gambar 9:

Pdosen-D1 = didapatkan dari CosD1 (Gambar 8) / CosD1 (Gambar 8 + jumlah kemunculan nama dosen pembimbing.

Pdosen-D2 = didapatkan dari CosD2 (Gambar 8) / CosD2 (Gambar 8) + jumlah kemunculan nama dosen pembimbing.

Dokumen	Nama Co pembimbing	Probabilitas Co Pembimbing (PCO)
PCOdosen-D1	Chaulina Alfianti Oktavia, S.Kom., M.T	0,007268074
PCOdosen-D2	-	-
PCOdosen-D3	Chaulina Alfianti Oktavia, S.Kom., M.T	0,076376883
PCOdosen-D4	Johan Ericka W P, M.Kom	0,07352347
PCOdosen-D5	Yekti Asmoro Kanthi, S.Si., M.A.B	0,25062916

Gambar 10. Perhitungan Cosine Similarity

Pdosen-D1 = Didapatkan dari CosD1 (Gambar 8) / CosD1 (Gambar 8 + jumlah kemunculan nama dosen pembimbing.

Pdosen-D2 = didapatkan dari CosD2 (Gambar 8) / CosD2(Gambar 8) + jumlah kemunculan nama dosen pembimbing.

Setelah nilai probabilitas pembimbing di temukan maka dilakukan perhitungan tingkat probabilitas pembimbing berdasarkan total kemunculan pada tiap data *training*.

No	Nama dosen	Tingkat Probabilitas Dosen Pembimbing (p)
P-1	Dr. Eva Handriyantini, S.Kom.,M.MT	0,1666667
P-2	Sugeng Widodo.,M.Kom	0,0833333
P-4	Diah Arifah P, S.Kom.,M.T	0,0833333
P-5	Anita, S.Kom, M.T.	0,0833333

Gambar 11. Perhitungan Probabilitas Dosen Berdasarkan Kategori.

Keterangan :

P-1 = Total kemunculan dosen / jumlah total dokumen

P-2 = Total kemunculan dosen / jumlah total dokumen.

No	Nama dosen	Tingkat P	Naive bayes
PCO-1	Chaulina Alfianti Oktavia, S.Kom., M.T	0,25	0,406563
PCO-2	Johan Ericka W P, M.Kom	0,1666667	0,301287
PCO-3	Yekti Asmoro Kanthi, S.Si., M.A.B	0,0833333	0,333962
PCO-4	Meivi Kartikasari, S.Kom., M.T	0,0833333	0,30138
PCO-5	Rakhmad Maulidi, M.Kom	0,1666667	0,358592

Gambar 12. Perhitungan Probabilitas Co dosen Pembimbing berdasarkan kategori.

Keterangan gambar 12:

PCO-1 = Total kemunculan dosen / jumlah total dokumen.

PCO-2 = Total kemunculan dosen / jumlah total dokumen.

Setelah semua variable yang di butuhkan sudah di dapatkan maka di lakukan perhitungan seperti pada rumus dibawah untuk dosen pembimbing sehingga akan didapatkan :

$$V_{Map} = \frac{\arg \max_{V_j \in V} P(x_1, x_2, x_3, \dots, x_n | V_j) P(V_j)}{P(x_1, x_2, x_3, \dots, x_n)} \dots (1)$$

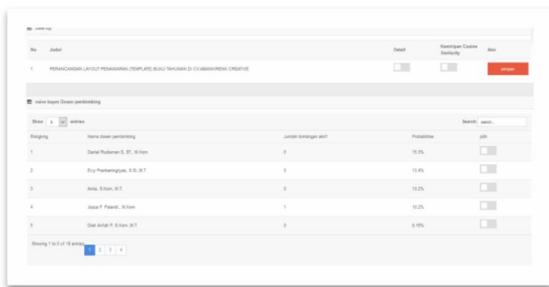
Keterangan :

V_j = Kategori data j = 1,2,3,...n

j₁ = Kategori Data1, dan j₂ = kategori Data2

No	Nama dosen	Naive bayes
NDP-1	Dr. Eva Handriyantini, S.Kom.,M.MT	0,185408
NDP-2	Sugeng Widodo.,M.Kom	0,200248
NDP-3	Diah Arifah P, S.Kom.,M.T	0,282101
NDP-4	Anita, S.Kom, M.T.	0,220309
NDP-5	Koko Wahyu Prasetyo, S.Kom.,M.T.I	0,333962

Gambar 13. Perhitungan Naive Bayes Dosen Pembimbing



Gambar 19. Halaman Detil Tentukan Dosen

Halaman Detil Validasi Data

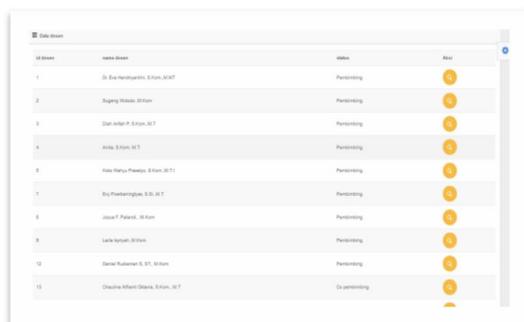
Pada halaman detil validasi data kaprodi dapat melihat detil form 7 yang sudah divalidasi atau yang sudah ditentukan pembimbing dan co dosen.



Gambar 20. Halaman Detil Validasi Data

Halaman Data Dosen

Pada halaman data dosen kaprodi dapat melihat daftar dosen pembimbing dan co pembimbing serta dapat melihat detil dosen dengan cara menekan tombol detil atau tombol bersimbol kaca pembesar.



Gambar 21. Halaman Data Dosen

4. KESIMPULAN DAN SARAN

Kesimpulan

Aplikasi Sistem penunjang keputusan penentuan dosen pembimbing menggunakan *Naive Bayes* dapat mempermudah kaprodi dalam menentukan dosen pembimbing dan co pembimbing dengan cara sistem dapat

mengeluarkan beberapa rekomendasi nama dosen pembimbing dan co pembimbing sebagai bahan penunjang keputusan yang didapatkan dari hasil perhitungan metode *Naive Bayes & Cosine Similarity*. Serta sistem dapat melihat kemiripan antar dokumen sebagai bahan pertimbangan persetujuan dan penentuan dosen pembimbing.

Saran

Berdasarkan penelitian yang telah dilakukan, membutuhkan beberapa syarat yang dapat dijadikan acuan pengembangan. Beberapa saran yang dapat dijadikan acuan untuk pengembangan aplikasi atau sistem lebih lanjut adalah Perlunya pengkajian lebih lanjut untuk penambahan data *training* agar akurasi atau tingkat keakuratan rekomendasi dosen pembimbing lebih akurat.

5. REFERENSI

anonymous, "informatikalogi.com," 2018. [Online]. Available : <https://informatikalogi.com/text-preprocessing/>. [Diakses 23 November 2018].

[2] D. & N. Vijayarani, "Preprocessing Techniques for Text Mining," 2015.

[3] S. pradnyana, "Perancangan Dan Implementasi Automated Document Integration Dengan Menggunakan Algoritma Complete Linkage Agglomerative Hierarchical Clustering.," 2012.

[4] S. H. Wijaya A. P, "Naive Bayes Classification pada Klasifikasi Dokumen Untuk Identifikasi Konten E-Government," 2016.

[5] Y. A.F, "Penerapan Metode Naive Bayes Dalam Pengklasifikasian Artikel Clickbait Berbahasa Indonesia.," 2017.

[6] P. Kasih dan I. N. Farid, "Sistem Bantu Pemilihan Dosen Pembimbing Tugas Akhir Berdasarkan Kateori Pilihan dan Keahlian Dosen menggunakan Naive Bayes," p. 62, 2017.