**USULAN TUGAS AKHIR**

1. **IDENTITAS PENGUSUL**

Nama : Muhammad Arizona Fauzie.

NRP : 5108100021.

Dosen Wali : Tohari Ahmad, S.Kom, MIT., Ph.D.

Dosen Pembimbing : 1. Diana Purwitasari, S.Kom., M.Sc.

2. Anny Yuniarti, S.Kom., MComp.Sc.

1. **JUDUL TUGAS AKHIR**

**“Identifikasi Profil Peneliti Berdasarkan Data Publikasi Dosen Menggunakan Algoritma K-Means++ dan Metode Analisis Jaringan Sosial”**

1. **LATAR BELAKANG**

Penelitian dalam tingkat perguruan tinggi harus berurusan dengan area ilmu pengetahuan yang semakin berkembang tetapi dengan sumber daya yang terbatas. Untuk itu pandangan yang jelas dan informasi mengenai aktifitas penelitian sangat penting untuk menciptakan institusi penelitian yang memungkinkan untuk mengetahui struktur kolaborasi antar peneliti sebagai cara yang efektif untuk meningkatkan kinerja ilmiah dan menambah efektifitas dan kualitas hasil penelitian [3]. Salah satu metode yang digunakan untuk mengetahui kolaborasi aktor adalah dengan menggunakan Analisis Jaringan Sosial. Analisis Jaringan Sosial merupakan kumpulan metode relasional yang digunakan untuk memahami dan mengidentifikasi relasi antar aktor secara sistematis. Pada Tugas Akhir ini dilakukan analisis terhadap struktur kolaborasi penelitian akademik dengan menggunakan Analisis Jaringan Sosial. Hal ini bertujuan untuk menciptakan pemahaman mendalam mengenai proses dimana aktor dalam Perguruan Tinggi berkomunikasi atau berhubungan satu dengan yang lainnya*.* Analisis Jaringan Sosialmemungkinkan untuk merepresentasikan kolaborasi aktor dalam perguruan tinggi ke dalam *graph*. Pada Analisis Jaringan Sosial, hubungan antar aktor didasarkan pada data berupa dokumen atau publikasi ilmiah.

Secara garis besar terdapat tiga tahapan utama pada Analisis Jaringan Sosial yaitu ekstraksi konsep, deteksi topik, dan analisis jaringan sosial [8]. Pada tahap ekstraksi konsep dilakukan ekstraksi terhadap abstraksi kata kunci data publikasi. Kata kunci tersebut kemudian dijadikan input pada *Latent Semantic Analysis* untuk mengidentifikasi konsep yang ada. Konsep yang telah dihasilkan kemudian dikelompokkan dengan menggunakan algoritma K-Means++. Pengelompokan menyebabkan konsep–konsep yang sama berkumpul menjadi sebuah topik. Tahap berikutnya adalah pembentukan *graph*. *Graph* memiliki dua elemen utama yaitu *node* dan *edge*. *Node* merupakan representasi aktor, sedangkan *edge* merupakan representasi relasi antar aktor. Bobot pada *edge* didasarkan pada kesamaan topik dari publikasi-publikasi yang dimiliki oleh peneliti. Tahap terakhir adalah menganalisis *graph* menggunakan Analisis Jaringan Sosial. Pada tahap ini dilakukan pengukuran nilai *degree*, *closeness*, dan *betweeness* dari masing-masing peneliti sehingga dapat diketahui kolaborasi antar peneliti. *Degree* digunakan untuk mengukur jumlah *edge* yang terhubung ke *node*. Peneliti dengan sentralitas *degree* yang tinggi adalah peneliti yang memiliki kolaborasi paling banyak. Dengan menggunakan pengukurann tersebut, dapat diidentifikasi peneliti yang paling aktif. *Closeness* didasarkan pada total jarak antara suatu *node* dengan *node* yang lain. Seorang peneliti dianggap memilki sentralitas *closeness* tinggi jika memiliki koneksi terpendek dengan peneliti lainnya. Hal ini berarti bahwa mereka cenderung berkolaborasi secara lebih mudah karena memiliki jarak terpendek dengan *node* lainnya. *Betweeness* mengukur kemampuan node untuk menghubungkan node yang tidak memiliki koneksi langsung. Node ini disebut dengan penghubung, karena memiliki kemampuan untuk menyalurkan informasi dari satu peneliti kepada peneliti lainnya.

Institusi pendidikan dengan bertahap perlu untuk menilai dan memperbaiki aktifitasnya agar dapat menyediakan keseimbangan antara aset yang nyata dan tidak, serta untuk mengukur kemampuan di masa depan maupun kinerja yang dulu. Analisis Jaringan Sosial secara efektif dapat digunakan untuk menganalisis hasil penelitian pada Perguruan Tinggi untuk identifikasi komunitas peneliti, peneliti yang paling aktif, peneliti yang menjadi penghubung, serta kekuatan jaringan peneliti yang menggambarkan kemampuan untuk mengembangkan ilmu pengetahuan. Analisis yang dilakukan terhadap data publikasi peneliti menyediakan rincian mengenai kolaborasi antara anggota institusi pendidikan.

1. **RUMUSAN MASALAH**

Rumusan masalah yang diangkat dalam Tugas Akhir ini adalah:

1. Bagaimana merepresentasikan kolaborasi peneliti akademis ke dalam *graph*?
2. Bagaimana mengidentifikasi peneliti yang paling aktif, peneliti dengan kolaborasi terbanyak, dan peneliti yang bertindak sebagai penghubung?
3. **BATASAN MASALAH**

Permasalahan yang dibahas dalam Tugas Akhir ini memiliki batasan sebagai berikut:

1. Data yang digunakan berasal dari database publikasi ITS. Setiap *record* data memiliki atribut:
   1. Judul Publikasi;
   2. NRP Mahasiswa;
   3. NIP Dosen Pembimbing 1;
   4. NIP Dosen Pembimbing 2;
   5. Abstraksi;
   6. Kata kunci.
2. *Library* yang digunakan antara lain:
   1. *Lucene* merupakan *library* sistem temu kembali informasi. Seluruh format dokumen kecuali citra dapat diindeks selama informasi tekstual yang dimiliki dapat diekstraksi.
   2. JUNG (Java *Universal Network Graph*) *Framework* merupakan *library* yang menyediakan peralatan untuk memodelkan, menganalisis, dan memvisualisasikan data yang dapat direpresentasikan sebagai *graph* atau jaringan.
3. Tugas Akhir ini dibuat menggunakan bahasa pemrograman Java.
4. **TUJUAN TUGAS AKHIR**

Tujuan Tugas Akhir ini adalah untuk membangun representasi *graph* dari profil peneliti yang berasal dari data publikasi dosen. Representasi *graph* tersebut kemudian dianalisis menggunakan analisa jaringan sosial untuk mengidentifikasi profil peneliti.

1. **MANFAAT TUGAS AKHIR**

Manfaat Tugas Akhir ini adalah untuk mengetahui kolaborasi antara peneliti dalam lingkungan akademis ITS, sehingga dapat meningkatkan kinerja ilmiah dan menambah efektifitas dan kualitas penelitian.

1. **TINJAUAN PUSTAKA**

Pada Tugas Akhir ini tahap pertama adalah melakukan ekstraksi konsep. Ekstraksi konsep dimulai dengan mengekstraksi kata kunci dari abstraksi data publikasi dosen yang ada dengan menggunakan metode *Rapid Automatic Keyphrase Extraction* (RAKE).

* 1. Ekstraksi kata kunci menggunakan *Rapid Automatic Keyphrase Extraction* (RAKE)

Metode RAKE merupakan metode *unsupervised*, serta tidak tergantung pada domain dan bahasa. Metode RAKE didasarkan pada pengamatan bahwa frase kunci seringkali terdiri dari beberapa kata tetapi jarang terdiri dari *stopword* [6].

Input parameter RAKE terdiri dari daftar *stopword* dan tanda baca sebagai pemisah teks menjadi kandidat kata kunci. Metode RAKE memiliki tahapan sebagai berikut [7]:

1. Ekstraksi kandidat.

Ekstraksi kandidat kata kunci dimulai dengan memisahkan teks menggunakan *stopword* dan tanda baca. Misalkan terdapat di adalah dokumen ke-i dari data publikasi yang ada. Setelah abstraksi diekstraksi kata kuncinya, maka didapatkan kandidat kata kunci *Tdi* = {*t1, t2, t3, ...,tn*}, dengan *t1, t2, t3, ...,tn* dalam bentuk kata, maupun frase kandidat kata kunci.

1. Menghitung matriks *co-ocurrence*.

Setelah kandidat kata kunci didapatkan, langkah selanjutnya adalah menghitung matriks *co-ocurrence.* Matriks *co-ocurrence* memetakan frekuensi kemunculan suatu kata dan frase kata kunci.

1. Menghitung rasio.

Nilai rasio merupakan perbandingan antara derajat kata dengan frekuensi kata. Derajat kata adalah jumlah kemunculan kata pada dokumen ditambah jumlah frase yang mengandung kata tersebut. Derajat kata pada matriks *co-ocurrence* didapat dari penjumlahan satu kolom atau satu baris. Frekuensi kata adalah jumlah kemunculan kata dalam teks. Nilai frekuensi bisa didapatkan pada nilai diagonal pada matriks *co-ocurrence*. Rasio dapat dihitung dengan menggunakan Persamaan 1.

(1)

dengan = derajat kata, dan

= frekuensi kata.

1. Menghitung nilai fitur dasar.

Nilai fitur dasar merupakan nilai penjumlahan rasio kata yang ada pada kandidat kata kunci. Nilai tersebut kemudian diurutkan berdasarkan nilai tertinggi sampai terendah. Nilai fitur dasar dihitung dengan menggunakan Persamaan 2.

N*fd =* rasiot + rasioft (2)

dengan N*fd* = nilai fitur dasar,

rasiot = rasio kata kunci *t*, dan

rasioft = rasio frase yang mengandung kata kunci *t*.

1. Pemilihan kandidat frase kunci dengan skor tertinggi.

Setelah kandidat kata kunci diberi nilai, selanjutnya dipilih sejumlah *k* kandidat dengan nilai tertinggi.

* 1. Ekstraksi konsep menggunakan *Latent Semantic Analysis* (LSA)

Kata kunci yang dihasilkan pada tahap ekstraksi kata kunci menggunakan *Rapid Automatic Keyphrase Extraction* (RAKE) digunakan sebagai data input untuk membentuk matriks *co-ocurrence* pada tahap ekstraksi konsep menggunakan *Latent Semantic Analysis* (LSA).

*Latent Semantic Analysis* (LSA) merupakan teori dan metode untuk mengekstraksi dan merepresentasikan arti penggunaan kontekstual kata menggunakan perhitungan statistik yang diterapkan pada kumpulan teks dalam jumlah besar [4]. LSA memiliki tahapan sebagai berkut:

1. Tahap pertama dalam LSA adalah merepresentasikan teks sebagai matriks dimana setiap baris mewakili kata yang unik dan setiap kolom mewakili bagian teks. Setiap sel berisi frekuensi kemunculan kata pada bagian teks. Setelah itu bentuk matriks *X* yang merupakan matriks dengan elemen (*i*, *j*) yang menunjukkan kemunculan term *i* pada dokumen *j*.
2. Buat matriks *XXT* yang menunjukkan *dot product* dari *term* antar dokumen.
3. Langkah selanjutnya adalah mengaplikasikan *Singular Value Decompotition* (SVD) pada matriks yang telah dibangun. SVD merupakan metode dekomposisi matriks yang membagi matriks menjadi dua kolom ortonormal, dan satu nilai singular pada diagonal matriks. Anggap U dan V merupakan matriks ortogonal dan ∑ merupakan matriks diagonal. Maka SVD dihitung dengan Persamaan 3.

X = U∑VT(3)

1. Tahap terakhir adalah menghitung similaritas pada ruang dimensi yang telah didekomposisi. Dilakukan dengan cara membandingkan ∑kdj dengan ∑kdq.
   1. Identifikasi topik menggunakan Algoritma K-Means++

Konsep-konsep yang dihasilkan pada tahap ekstraksi konsep menggunakan *Latent Semantic Analysis* (LSA) digunakan sebagai input pada tahap identifikasi topik menggunakan K-Means++.

Pada *data mining* K-Means++ merupakan algoritma pemilihan nilai awal untuk algoritma K-Means. Metode ini diusulkan pada tahun 2007 oleh David Arthur dan Sergei Vassilvitski sebagai algoritma pendekatan untuk permasalahan *NP-hard* K-Means yang merupakan cara untuk menghindari *clustering* yang buruk yang dapat dihasilkan dari algoritma K-Means standar. Permasalahan K-Means adalah dalam mencari pusat *cluster* untuk meminimalisir variansi didalam kelas. Walaupun mencari solusi yang sesungguhnya dalam permasalahan k-means untuk input yang ada merupakan *NP-hard*, pendekatan standar untuk mencari solusi digunakan secara luas dan seringkali menemui solusi yang masuk akal dengan cepat. Bagaimanapun juga, algoritma K-Means memiliki dua teori penghalang utama yaitu:

1. Telah ditunjukkan bahwa waktu eksekusi terburuk dari algoritma ini adalah *super-polynomial* sebanding dengan ukuran input.
2. Pendekatan yang ditemukan dapat menjadi buruk dengan harapan pada fungsi objektif jika dibandingkan dengan *cluster* yang optimal.

Algoritma K-Means++ ditujukan pada penghalang kedua dengan mengutamakan tahap inisialisasi pusat *cluster* sebelum memproses lebih lanjut menggunakan iterasi algoritma K-Means standar. Dengan inisialisasi K-Means++, algroitma dijamin mendapatkan solusi kompetitif O(log k) terhadap solusi optimal K-Means.

Algoritma K-Means++ memiliki tahapan sebagai berikut [2]:

1. Pilih sebuah pusat secara seragam dengan acak dari seluruh titik data yang ada.
2. Untuk setiap titik data *x*, hitung nilai *D(x)*, dengan *D(x)* merupakan jarak antara titik data *x* dan pusat terdekat yang telah dipilih.
3. Pilih satu titik data baru secara acak sebagai pusat baru dengan menggunakan distribusi probabilitas terbobot dimana sebuah titik *x* dipilih menggunakan probabilitas sebanding dengan .
4. Ulangi langkah 2 dan 3 sampai terpilih pusat sebanyak *k*.
5. Setelah pusat awal dipilih, lanjutkan dengan menggunakan algoritma K-Means standar.

Algoritma K-Means standar sendiri memiliki tahapan sebagai berikut:

1. Pilih k pusat awal secara acak *C* ={*c1, c2, c3, ..., ck*}*.*
2. Untuk setiap i {1, ..., *k*}, jadikan *cluster* *C1* sebagai kumpulan titik di *X* yang lebih dekat dengan *ci* daripada *cj* untuk semua *j* ≠ *i*.
3. Untuk setiap i {1, ..., *k*}, jadikan ci sebagai pusat dari seluruh titik pada *Ci* sehingga *ci* = .
4. Ulangi tahap 2 dan 3 sampai semua pusat tidak berubah.
   1. Analisis jaringan menggunakan metode Analisis Jaringan Sosial

Jaringan sosialmerupakan konstruksi teoritis yang digunakan dalam ilmu pengetahuan sosial untuk mempelajari relasi antar individu, group, organisasi, maupun keseluruhan masyarakat. Kompleksitas relasi yang ada pada jaringan sosial tidak memungkinkan untuk mengidentifikasi relasi sosial hanya dengan menggunakan pemahaman intuitif. Untuk itu diperlukan Analisis Jaringan Sosialyang menyediakan metode secara empiris untuk memperluas pemahaman intuitif mengenai pola–pola yang membentuk struktur sosial.

Analisis Jaringan Sosialmerupakaan kumpulan metode relasional yang digunakan untuk memahami dan mengidentifikasi relasi antar aktor secara sistematis [5]. Analisis Jaringan Sosial pada umumnya:

1. Didorong oleh intuisi struktural berdasarkan relasi yang menghubungkan aktor sosial;
2. Didasarkan pada data empiris;
3. Digambarkan pada citra grafis;
4. Bergantung pada penggunaan model perhitungan.

Bagian terpenting dalam jaringan sosial adalah perpaduan antara aktor dan relasinya. Aktor direpresentasikan sebagai titik, *node* atau *vertex*. Relasi direpresentasikan dengan garis atau *edge*. Pada umunya relasi dapat memiliki nilai biner atau nilai tertentu, serta memiliki arah maupun tidak berarah.

Data pada Analisis Jaringan Sosial memuat setiap relasi antar *node*. Sumber data meliputi survei, akun yang sudah terpublikasi, informasi khusus, dan lain sebagainya.

Bentuk atau topologi jaringan dapat mempengaruhi aliran pada jaringan. Terdapat dua fitur penting dalam topologi jaringan yaitu konektifitas dan sentralitas. Konektifitas merujuk pada bagaimana aktor pada suatu bagian jaringan terhubung dengan aktor pada jaringan yang lainnya. Sentralitas merujuk pada lokasi, identifikasi dimana sebuah aktor berada pada suatu jaringan. Terdapat tiga fitur pada konektifitas jaringan yaitu:

1. *Reachability.*

Kemungkinan sebuah aktor *i* untuk menjangkau aktor *j*. Kondisi ini bernilai benar jika terdapat relasi yang menghubungkan antara suatu aktor dengan aktor yang lainnya.

1. *Distance.*

Jika suatu aktor dapat dijangkau maka berapa langkah untuk mencapai aktor tersebut.

1. *Number of path.*

Berapa jumlah jalan yang berbeda yang menghubungkan kedua aktor tersebut.

Pada level individu, posisinya dalam jaringan dapat diketahui melalui sentralitas. Secara konseptual, sentralitas mengidentifikasi *node* yang menjadi pusat jaringan, sehingga dapat diketahui node yang penting. Terdapat tiga pengukuran sentralitas secara umum, yaitu [1]:

1. *Degree.*

*Degree* merupakan jumlah relasi. Aktor dengan relasi terbanyak adalah aktor yang paling penting. Dapat dihitung menggunakan Persamaan 4.

dengan = *centrality degree,*

= *degree* dari aktor (*node*) *i*, dan

= *edge* antara aktor *i* dan *j*.

1. *Closeness.*

*Closeness* merupakan kedekatan suatu aktor dengan aktor yang lainnya. Aktor dikatakan penting apabila dekat dengan aktor yang lain. *Closeness* didasarkan pada kebalikan jarak dari tiap aktor ke aktor yang lain pada jaringan. Dapat dihitung menggunakan Persamaan 5.

dengan = *centrality* *closeness* aktor (*node*) *i* dan

= jarak aktor *i* ke *j*.

Rumus tersebut dapat dinormalisasi menjadi Persamaan 6.

dengan = *normalized* *centrality closeness* aktor (*node*) *i* dan

g = jumlah aktor.

1. *Betweeness.*

*Betweeness* menghitung jumlah jalan terpendek antara aktor j dan k dimana aktor i berada. Dapat dihitung menggunakan Persamaan 7.

dengan = c*entrality* *betweeness* aktor (*node*) *i*,

gjk (ni) = jumlah *path* dimana aktor *i* berada, dan

gjk = jumlah *path* yang menghubungkan aktor *j* dan *k*.

Rumus tersebut dapat dinormalisasi menjadi Persamaan 8.

dengan = *normalized centrality betweeness* aktor (*node*) *i* dan

g = jumlah aktor.

1. **RINGKASAN TUGAS AKHIR**

Secara garis besar terdapat tiga tahapan utama pada Analisis Jaringan Sosial yaitu ekstraksi konsep, deteksi topik, dan analisis jaringan sosial [8]. Pada tahap ekstraksi konsep dilakukan ekstraksi terhadap abstraksi kata kunci data publikasi. Kata kunci tersebut disimpan ke dalam database dan kemudian dijadikan input pada *Latent Semantic Analysis* untuk mengidentifikasi konsep yang ada. Konsep yang telah dihasilkan kemudian dikelompokkan dengan menggunakan algoritma K-Means. Pengelompokan menyebabkan konsep–konsep yang sama berkumpul menjadi sebuah topik. Tahap berikutnya adalah pembentukan *graph*. *Graph* memiliki dua elemen utama yaitu *node* dan *edge*. *Node* menrupakan representasi aktor, sedangkan *edge* merupakan representasi relasi antar aktor. Relasi antar aktor didasarkan pada kesamaan topik dari publikasi-publikasi yang dimiliki oleh peneliti. Dokumen-dokumen yang dimiliki oleh peneliti memiliki nilai similaritas terhadap masing-masing topik yang ada. Terdapat peneliti yang memilki topik publikasi yang sama sehingga nantinya dapat diketahui terdapat relasi atau tidak antar peneliti tersebut. Tahap terakhir adalah menganalisis *graph* menggunakan Analisis Jaringan Sosial. Pada tahap ini dilakukan pengukuran nilai *degree*, *closeness*, dan *betweeness* dari masing-masing peneliti sehingga dapat diketahui kolaborasi antar peneliti. *Degree* digunakan untuk mengukur jumlah *edge* yang terhubung ke *node*. Peneliti dengan sentralitas *degree* yang tinggi adalah peneliti yang memiliki kolaborasi paling banyak. Dengan menggunakan pengukuran tersebut, dapat diidentifikasi peneliti yang paling aktif. *Closeness* didasarkan pada total jarak antara suatu *node* dengan *node* yang lain. Seorang peneliti dianggap memilki sentralitas *closeness* tinggi jika memilik koneksi terpendek dengan peneliti lainnya. Hal ini berarti bahwa mereka cenderung berkolaborasi secara lebih mudah karena memiliki jarak terpendek dengan *node* lainnya. *Betweeness* mengukur kemampuan *node* untuk menghubungkan *node* yang tidak memiliki koneksi langsung. Node ini disebut dengan penghubung, karena memiliki kemampuan untuk menyalurkan informasi dari satu peneliti kepada peneliti lainnya. Diagram alur mengenai Tugas Akhir ini dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alur Identifikasi Profil Peneliti Berdasarkan Data Publikasi Dosen Menggunakan Algoritma K-Means++ dan Analisis Jaringan Sosial.

1. **METODOLOGI**

Metodologi yang dilakukan dalam Tugas Akhir ini memiliki beberapa tahapan sebagai berikut :

1. Penyusunan Proposal Tugas Akhir

Tahap awal untuk memulai pengerjaan Tugas Akhir adalah dengan menyusun Proposal Tugas Akhir. Pada proposal ini penulis mengajukan gagasan mengenai “Identifikasi Profil Peneliti Berdasarkan Data Publikasi Dosen Menggunakan Algoritma K-Means dan Metode Analisis Jaringan Sosial”.

1. Studi Literatur

Pada tahap ini dilakukan pencarian, pengumpulan, penyaringan, pembelajaran, dan pemahaman literatur mengenai ekstraksi kata kunci menggunakan metode *Rapid Automatic Keyphrase Extraction* (RAKE), ekstraksi konsep menggunakan *Latent Semantic Analysis* (LSA), pengelompokan konsep menggunakan K-Means++, dan Analisis Jaringan Sosial. Literatur pada Tugas Akhir ini sebagian besar berasal dari internet berupa artikel ilmiah, makalah, materi kuliah, serta buku–buku referensi.

1. Implementasi

Implementasi merupakan tahap membangun sistem identifikasi profil peneliti berdasarkan data publikasi dosen menggunakan K-Means++ dan metode Analisis Jaringan Sosial.

1. Pengujian dan Evaluasi

Pada tahap ini dilakukan uji coba terhadap sistem yang telah dibuat meliputi pengamatan kinerja sistem dan kendala yang mungkin timbul.

1. Penyusunan Buku Tugas Akhir

Tahap terakhir merupakan penyusunan laporan yang memuat dokumentasi mengenai pembuatan sistem serta hasil dari implementasi sistem yang telah dibuat. Bagian inti dari buku Tugas Akhir ini terdiri dari :

1. Pendahuluan
   1. Latar Belakang Masalah
   2. Rumusan Masalah
   3. Batasan Masalah
   4. Tujuan
   5. Manfaat
2. Tinjauan Pustaka
   1. Hasil penelitian terdahulu yang relevan
   2. Teori yang digunakan untuk menyelesaikan permasalahan
3. Metodologi
   1. Metode yang digunakan
   2. Bahan dan peralatan yang digunakan
   3. Urutan pelaksanaan percobaan
4. Hasil dan Pembahasan
   1. Data hasil percobaan atau pengukuran
   2. Pembahasan (analisis, sintesis, dan evaluasi)
5. Kesimpulan dan Saran
6. **JADWAL PEMBUATAN TUGAS AKHIR**

Jadwal pembuatan Tugas Akhir dapat dilihat pada Tabel 1. di bawah ini.

Tabel 1. Jadwal Pembuatan Tugas Akhir

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Tahapan** | **Bulan** | | | | | | | | | | | | | | | | | | | |
| **Maret** | | | | **April** | | | | **Mei** | | | | **Juni** | | | | **Juli** | | | |
| 1 | Pembuatan Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Implementasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Pengujian dan Evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Penyusunan Buku Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. **DAFTAR PUSTAKA**

[1] Anastasios, T., Cleo S., Effic P., Olivier T, & Georgios M. (2012). *”Co-authorship Networks in Academis Researh Commmunities: The Role of Network Strengh”*. Technological Educational Institute of Athens.

[2] Arthur, D., & Sergei V. (2007). *“K-means++: The Advantage of Carefull Seeding”*. Proceedings of The Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics Philadelphia, PA, USA. Pp. 1027-1035.

[3] Bento, C, & Hideaki T. (2013). *“Finding Research Communities and Their Relationship by Anayzing the Co-Authorship Network”.* National Institute of Informatics, Japan.

[4] Landauer, T. K., Foltz, P. W., & Laham, D. (1998*). “Introduction to Latent Semantic Analysis”*, Discourse Processes, 25, 259-284.

[5] Moody, J. (2004). “*Social Network Analysis*”. American Sociological Association, San Francisco.

[6] Rose, S., Engel, D., Cramer, N., & Cowley, W. (2010). *“Automatic Keyword Extraction from Individual Documents”*. Text Mining: Applications and Theory.

[7] Ulinnuha, N., R.V. Hari G., Diana P. (2013*).“Ekstraksi Kata Kunci Menggunakan M-RAKE pada Dokumen Penelitian”*. Teknik Informatika, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.

[8] Velardi, P., Roberto N., Alessandro C., & Fulvio D. (2013). *“A New Content-Based Model for Social Network Analysis”*. University of Roma, & University of Marche.