

## BAB IV

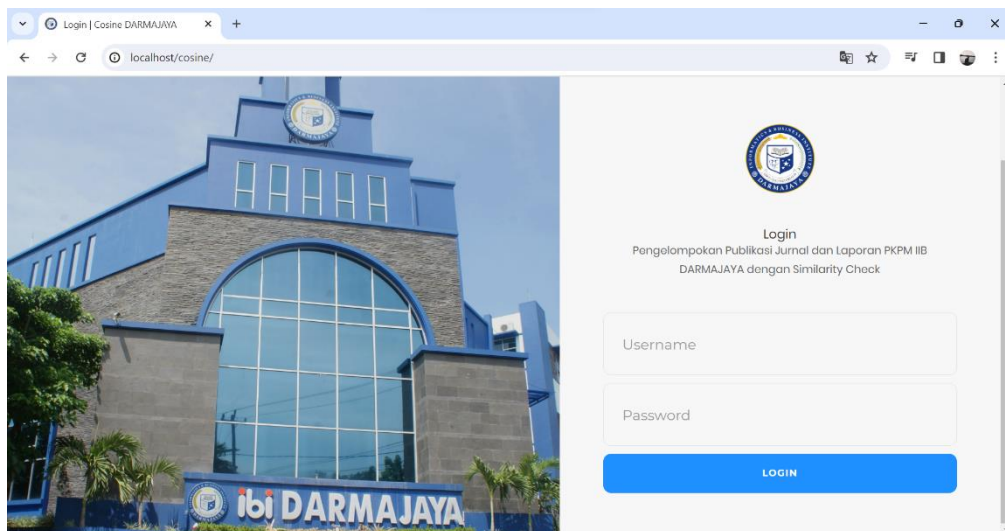
### HASIL PENELITIAN DAN PEMBAHASAN

#### 4.1 Hasil Penelitian

Hasil penelitian yang dilakukan di LPPM berupa data wawancara serta observasi terkait cara upload, penyimpanan, data, serta menentukan tema dan pengelompokan. Data yang di dapat kemudian diimplementasikan kedalam *Website* pengelompokan publikasi jurnal dan laporan PKPM IIB Darmajaya. *Website* ini dapat dibuka di perangkat komputer maupun *smartphone* yang dapat mengakses internet. Adapun alamatnya dapat diakses pada “localhost/cosine”.

##### 4.1.1 Tampilan *Log In*

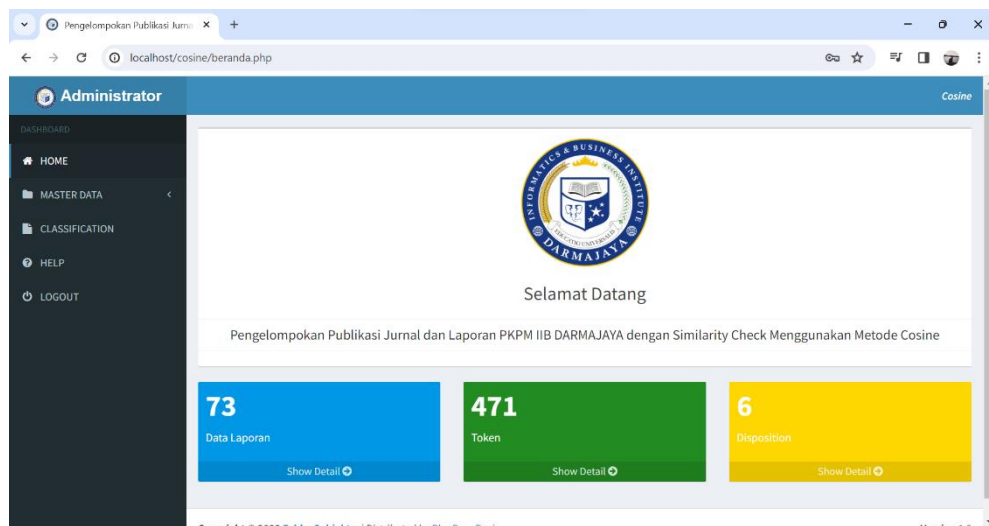
Fitur *login* akan menampilkan menu untuk masuk kedalam sistem dengan memasukkan *username* dan *password* kemudian klik *login*. Menu ini digunakan oleh pengguna untuk masuk kedalam menu utama yang berfungsi untuk keamanan pada sistem sehingga tidak semua pengguna dapat masuk kedalam sistem, dan juga ketika mahasiswa ingin mengunggah laporan ataupun jurnal, data dari mahasiswa sudah teridentifikasi sehingga mahasiswa tidak perlu lagi mengisi form ketika ingin mengunggah laporan dan jurnal. Gambar 4.1 berikut merupakan tampilan interface *Log in* :



**Gambar 4.1 Tampilan *Log In***

#### 4.1.2 Tampilan Menu Utama

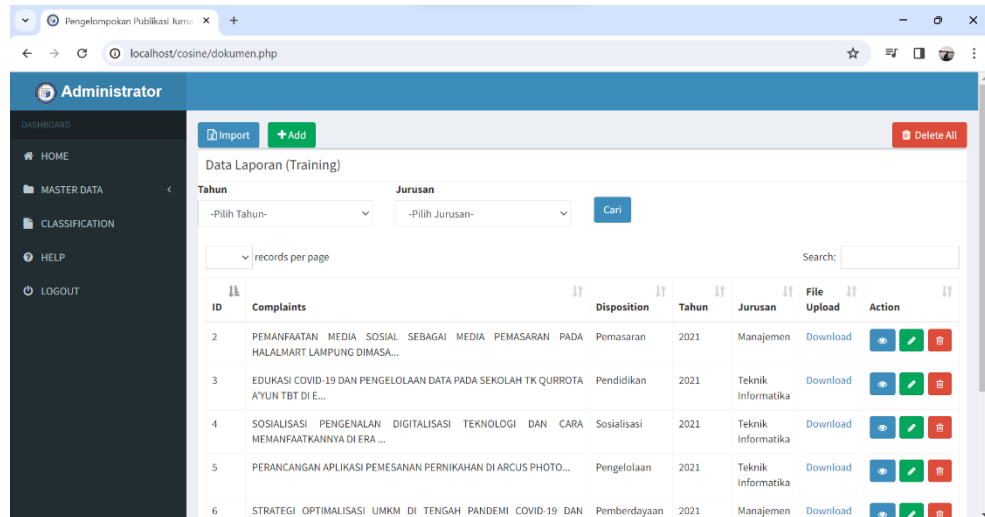
Menu utama pada halaman utama ini ini terdapat beberapa tampilan menu yaitu menu *home* yaitu menu tampilan utama, menu master data yang berisikan inputan data *pra-prosesing* sebelum melakukan perhitungan *cosine* yaitu data akan dilakukan pembagian data *training* selanjutnya akan ada tahapan *case folding*, *tokenizing*, *weigth value*, *basic word*, dan *dispostion list*, setelah semua tahapan selesai maka akan masuk menu *classificarion*, menu *help* untuk informasi bantuan sistem dan *logout* untuk keluar dari sistem. Gambar 4.2 berikut merupakan tampilan Menu Utama :



Gambar 4.2 Tampilan Menu Utama

#### 4.1.3 Tampilan Data Laporan (*Training*)

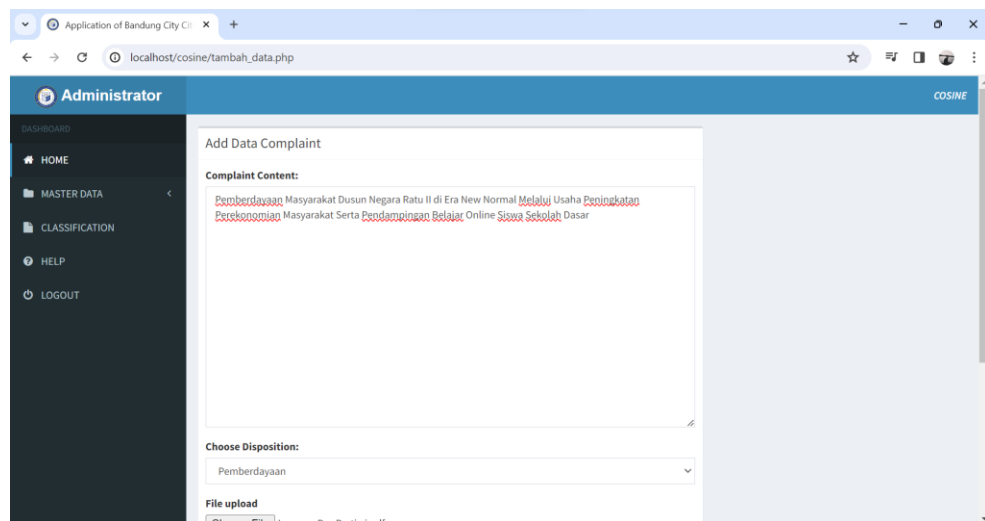
Menu data laporan (*training*) merupakan menu untuk mengelola data laporan (*training*) yang akan dilakukan perhitungan, pada menu ini disediakan untuk menjadi bahan pembelajaran model agar model dapat menggeneralisasi (menemukan pola) data sehingga nantinya dapat digunakan untuk memprediksi data baru sebelum dilakukan dalam perhitungan. Gambar 4.3 berikut merupakan tampilan menu Data Laporan :



**Gambar 4.3 Tampilan menu Data Laporan (Training)**

#### 4.1.4 Tampilan Penguploadan

Mahasiswa diminta untuk memasukan judul laporan, mengupload laporan, tahun dilaksanakan PKPM, dan pilih jurusan. Guna untuk mengklasifikasi kan laporan mahasiswa tersebut masuk ke kategori tema yang akan di baca oleh system menggunakan *cosine similarity*. Gambar 4.4 berikut merupakan tampilan penguploadan :

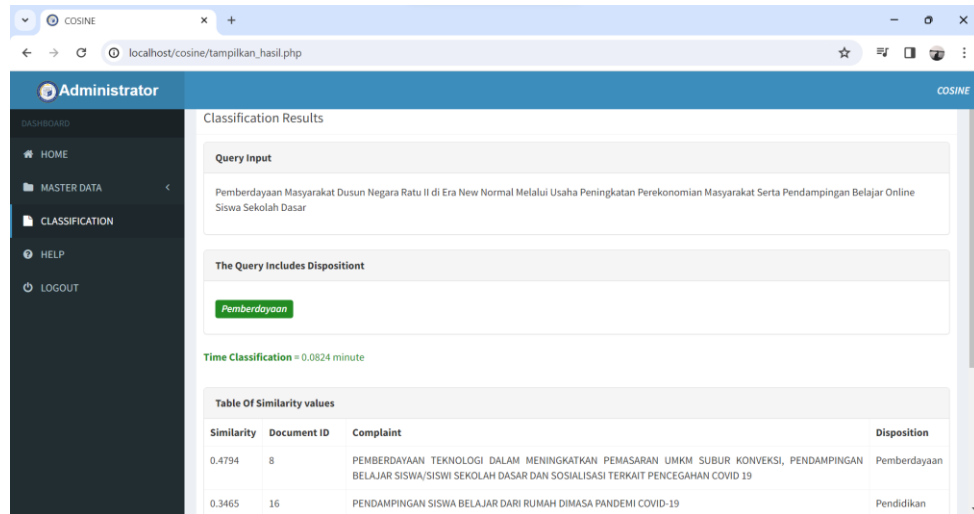


**Gambar 4.4 Tampilan Penguploadan**

#### 4.1.5 Menu Clasification Results

Setelah melakukan penguploadan mahasiswa akan mengetahui berapa nilai dari Time Classification, serta dapat mengetahui pengelompokan tema (Disposition) dari judul laporan mereka, dan dapat melihat data-data terdahulu.

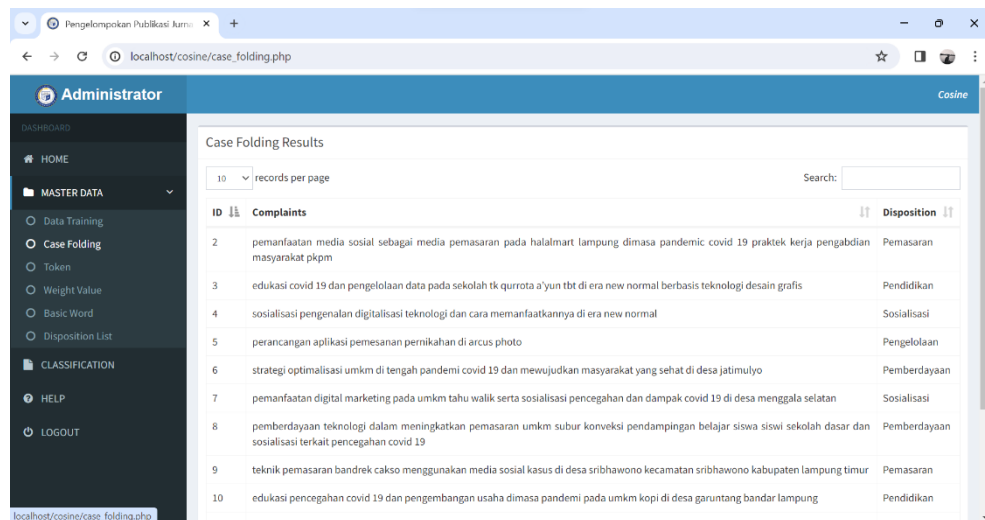
Gambar 4.5 merupakan tampilan *Classification Results*:



**Gambar 4.5 Tampilan *Classification Results***

#### 4.1.6 Tampilan *Case Folding*

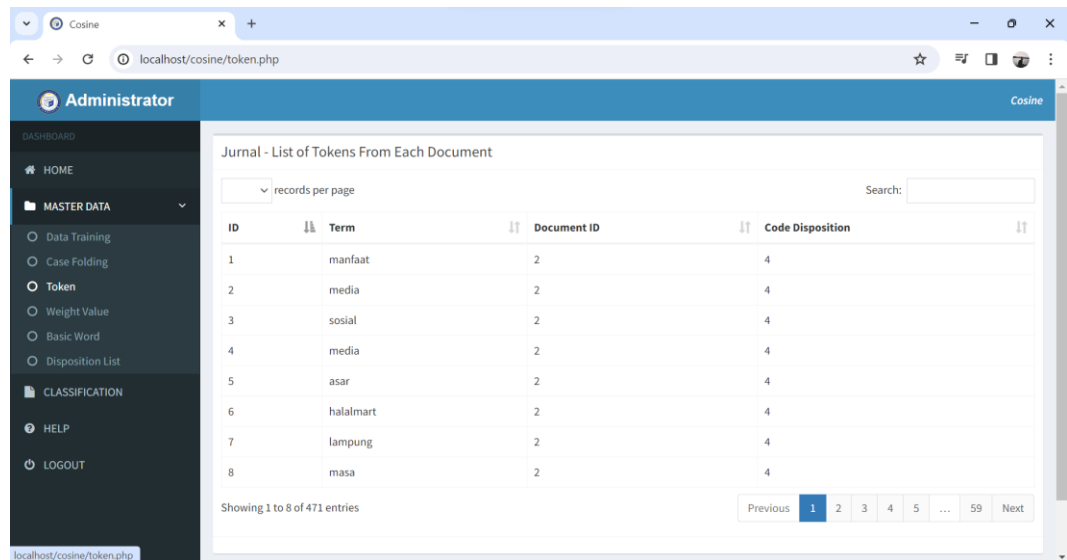
Menu *case folding* adalah menu untuk mengelola data *case folding* yang berfungsi untuk proses dalam *text preprocessing* yang dilakukan untuk menyeragamkan karakter pada data seperti mengubah seluruh huruf menjadi huruf kecil. Pada proses ini karakter-karakter 'A'-'Z' yang terdapat pada data diubah kedalam karakter 'a'-'z'. Gambar 4.6 merupakan tampilan *Case Folding* :



**Gambar 4.6 Tampilan *Case Folding***

#### 4.1.7 Tampilan Token

Tampilan menu token merupakan menu untuk mengelola data tokenizing dalam proses datamining dengan proses pemotongan string input berdasarkan tiap kata yang menyusunnya. Pemecahan kalimat menjadi kata-kata tunggal dilakukan dengan men-scan kalimat dengan pemisah white space (spasi, tab, dan newline). Gambar 4.7 merupakan tampilan menu Token :



| ID | Term      | Document ID | Code Disposition |
|----|-----------|-------------|------------------|
| 1  | manfaat   | 2           | 4                |
| 2  | media     | 2           | 4                |
| 3  | sosial    | 2           | 4                |
| 4  | media     | 2           | 4                |
| 5  | asar      | 2           | 4                |
| 6  | halalmart | 2           | 4                |
| 7  | lampung   | 2           | 4                |
| 8  | masa      | 2           | 4                |

Gambar 4.7 Tampilan Menu Token

#### 4.1.8 Tampilan Weigh Value

Tampilan *weigh value* merupakan menu untuk mengelola data *weigh value* dengan penentuan setiap variabel kata yang akan dilakukan proses *datamining* pada proses ini setiap kata akan dinilai untuk dapat mengetahui hasil kata yang dinilai. Gambar 4.8 merupakan tampilan *Weigh Value* :

Weight TF-IDF Value List From Each Document

| ID | Term        | Document ID | Count | Weight TF_IDF |
|----|-------------|-------------|-------|---------------|
| 1  | manfaat     | 2           | 1     | 0.8346        |
| 2  | jual        | 28          | 2     | 2.0214        |
| 3  | media       | 2           | 2     | 1.5354        |
| 4  | sosial      | 2           | 1     | 0.8346        |
| 5  | ukm         | 28          | 2     | 2.6235        |
| 6  | martabak    | 28          | 2     | 3.2256        |
| 7  | tradisional | 28          | 2     | 3.2256        |
| 8  | asli        | 28          | 2     | 3.2256        |

Showing 1 to 8 of 525 entries

**Gambar 4.8 Tampilan Weigh Value**

#### 4.1.9 Tampilan *Basic Word*

Tampilan *basic word* merupakan menu untuk mengubah kata menjadi kata kunci yang akan dilakukan proses *datamining* sehingga dapat di proses perhitungan untuk kalimat yang sama. Gambar 4.9 merupakan tampilan *Basic Word* :

Basic Word List

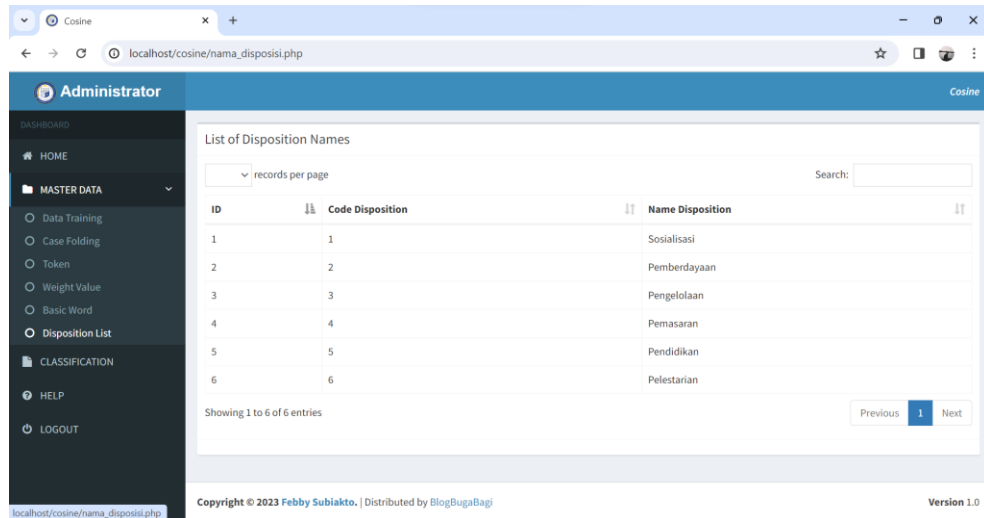
| ID | Basic Word | Type Basic Word |
|----|------------|-----------------|
| 1  | a          | Nomina          |
| 2  | ab         | Nomina          |
| 3  | aba        | Nomina          |
| 4  | aba-aba    | Nomina          |
| 5  | abad       | Nomina          |
| 6  | abadi      | Adjektiva       |
| 7  | abadiah    | Nomina          |
| 8  | abah       | Nomina          |
| 9  | abai       | Adjektiva       |
| 10 | abaimana   | Nomina          |

**Gambar 4.9 Tampilan Basic Word**

#### 4.1.10 Tampilan *Disposition List*

Menu *disposition list* merupakan menu untuk pengelompokan kata sebelum terjadi perhitungan *cosine* pada menu ini akan diberikan kode dan kata yang akan di proses *datamining*. Gambar 4.10 merupakan tampilan *Disposition*

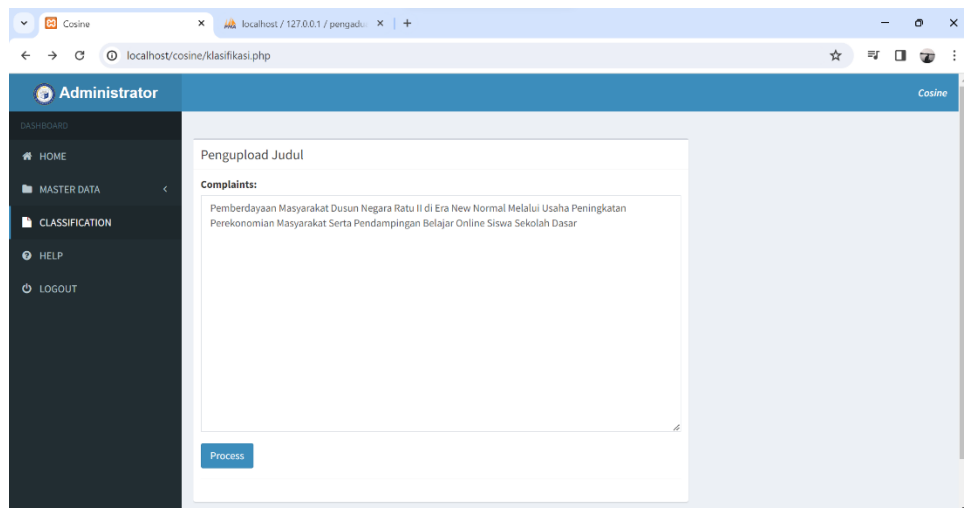
List :



Gambar 4.10 Disposition List

#### 4.1.11 Tampilan Clasification

Menu *clasification* merupakan menu untuk mengetahui data laporan yang akan di unggah akan masuk di kategori apa, tanpa harus mengunggah laporan. Pada menu ini terdapat tombol proses yang dimana *user* di minta untuk memasukkan judul untuk memproses data yang akan dilakukan perhitungan teknik *cosine*. Gambar 4.11 merupakan tampilan *Clasification*.

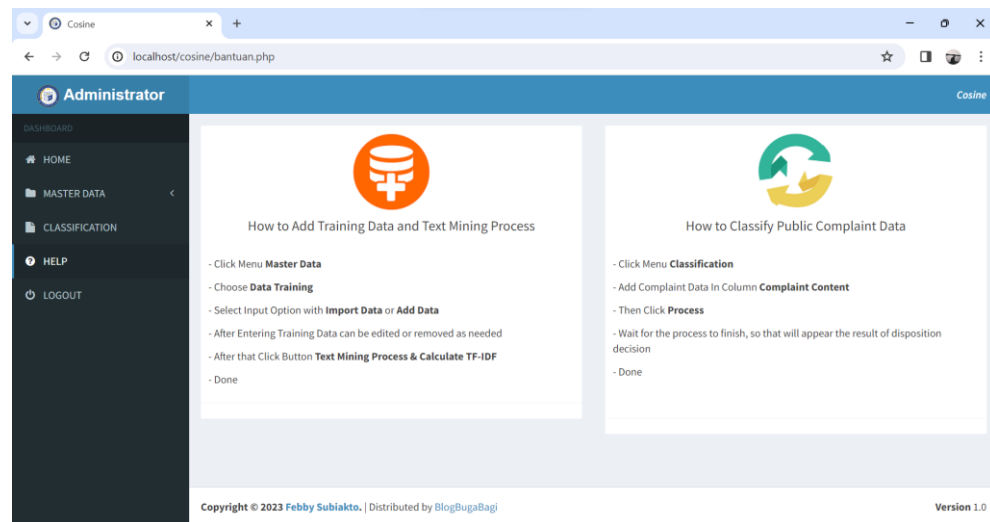


Gambar 4.11 Tampilan Clasification

#### 4.1.12 Tampilan Help

Menu *help* merupakan menu untuk melihat bantuan dalam aplikasi. Pada

menu ini terdapat informasi lengkap mengenai alur kegiatan proses *cosine*. Gambar 4.12 merupakan tampilan *Help* :



**Gambar 4.12 Tampilan *Help***

## 4.2 Pembahasan

Proses pengelompokan publikasi jurnal dan laporan PKPM IIB Darmajaya di implementasikan dalam bentuk website guna memudahkan mahasiswa serta pihak LPPM dalam mengetahui pengelompokan tema pada data laporan. Proses pengelompokan menggunakan metode *Cosine Similarity*, yang dimana dilakukan perhitungan *Cosine* agar menentukan hasil terbaik dalam suatu pencarian.

### 4.2.1 Perhitungan Algoritma *Cosine*

tahapan pertama melakukan *stemming* dari data yang ada pada database proses *stemming* menemukan kata dasar dengan menggunakan *stopword*. *Stopword* merupakan kata umum yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna, seperti “yang”, “di”, “ke: dan lain lain. Proses *stemming* diawali dengan melakukan proses menghilangkan imbuhan, awalan dan akhiran dari proses tersebut kemudian ditemukan kata dasar dari suatu kata di dalam dokumen yang di masukkan dalam kata kunci pencarian ataupun kata kunci dari judul dokumen. Berikut merupakan sampel data dalam melakukan proses *stemming* ditunjukkan pada tabel 3.1



**Tabel 4. 1** Data set Arsip pada database

| Dokumen | Term yang mewakili dokumen   |
|---------|--|
| Q       | Data keterangan mentah belum diolah belum bisa di-jaikan acuan mengambil suatu keputusan |
| D1      | Data-data mentah sekumpulan  |
| D2      | Belum bisa data diolah fakta informasi mentah menyajikan                                 |

Selanjutnya, proses yang dilakukan yaitu Indeksasi/Tokenisasi dan Hitung *TF* (*Term Frequency*), *DF* (*Document Frequency*), dan *IDF* (*Invers Document Frequency*)

1. Pisahkan tiap kata
2. Kata yang sama dianggap menjadi 1 kata
3. *TF* : Hitung jumlah kata yang muncul pada Dokumen
4. *DF* : Jumlah Seluruh *TF*
5. *IDF* : *Invert DF*

**Tabel 4. 2** Data Hasil tokenisasi

| No | Term       | TF |    |    | Df | Idf             |
|----|------------|----|----|----|----|-----------------|
|    |            | Q  | D1 | D2 |    | Log<br>(n/df)+1 |
| 1  | Data       | 1  | 1  | 1  | 3  | 1,425969        |
| 2  | Mentah     | 1  | 1  | 1  | 3  | 1,425969        |
| 3  | Sekumpulan | 0  | 1  | 0  | 1  | 1,90309         |
| 4  | Belum      | 1  | 0  | 1  | 2  | 1,60206         |
| 5  | Diolah     | 1  | 0  | 1  | 2  | 1,60206         |
| 6  | Fakta      | 0  | 0  | 1  | 1  | 1,90309         |
| 7  | menyajikan | 0  | 0  | 1  | 1  | 1,90309         |
| 8  | Acuan      | 1  | 0  | 0  | 1  | 1,90309         |

Setelah ditemukan nilai *TF* (*Term Frequency*), *DF* (*Document Frequency*), dan *IDF* (*Invers Document Frequency*) pada setiap term, maka selanjutnya dilakukan perhitungan bobot *wdt* (*weight document term*) dari setiap term yang ada.

**Tabel 4. 3** Tabel hasil perhitungan *wdt* (*weight document term*)

|   |            | Q        | D1       | D2       |
|---|------------|----------|----------|----------|
| 1 | Data       | 1,425969 | 1,425969 | 1,425969 |
| 2 | Mentah     | 1,425969 | 1,425969 | 1,425969 |
| 3 | Sekumpulan | 0        | 1,90309  | 0        |

|   |            | Q       | D1 | D2      |
|---|------------|---------|----|---------|
| 4 | Belum      | 1,60206 | 0  | 1,60206 |
| 5 | Diolah     | 1,60206 | 0  | 1,60206 |
| 6 | Fakta      | 0       | 0  | 1,90309 |
| 7 | menyajikan | 0       | 0  | 1,90309 |
| 8 | Acuan      | 1,90309 | 0  | 0       |

Setelah dilakukan perhitungan *wdt* (*weight document term*) dengan hasil masing masing term dokumen memiliki bobot tersendiri, selanjutnya perlu dilakukan perhitungan perkalian vektor.

**Tabel 4. 4** Perkalian vektor Q dan D

| No          | Tream      | Q          | D1        | D2        |
|-------------|------------|------------|-----------|-----------|
| 1           | Data       | 2,03338759 | 2,0333876 | 2,0333876 |
| 2           | Mentah     | 2,03338759 | 2,0333876 | 2,0333876 |
| 3           | Sekumpulan | 0          | 0         | 0         |
| 4           | Belum      | 0          | 0         | 0         |
| 5           | Diolah     | 1,60206    | 1,60206   | 1,60206   |
| 6           | Fakta      | 0          | 0         | 0         |
| 7           | menyajikan | 0          | 0         | 0         |
| 8           | Acuan      | 1,90309    | 0         | 0         |
| Total D0*D1 |            | 7,5719252  | 5,668835  | 5,668835  |
| Rata-rata   |            | 2,7517131  | 2,380932  | 2,380932  |

Telah ditentukan tingkat kemiripan dokumen yang ditunjukkan pada tabel dibawah ini.

**Tabel 4. 5** Hasil perhitungan *cosine similarity*

| No          | Trem       | Q          | D1        | D2        |
|-------------|------------|------------|-----------|-----------|
| 1           | Data       | 2,03338759 | 2,0333876 | 2,0333876 |
| 2           | Mentah     | 2,03338759 | 2,0333876 | 2,0333876 |
| 3           | Sekumpulan | 0          | 0         | 0         |
| 4           | Belum      | 0          | 0         | 0         |
| 5           | Diolah     | 1,60206    | 1,60206   | 1,60206   |
| 6           | Fakta      | 0          | 0         | 0         |
| 7           | menyajikan | 0          | 0         | 0         |
| 8           | Acuan      | 1,90309    | 0         | 0         |
| Total D0*D1 |            | 7,5719252  | 5,668835  | 5,668835  |

|           |      |           |          |          |
|-----------|------|-----------|----------|----------|
| No        | Trem | Q         | D1       | D2       |
| Rata-rata |      | 2,7517131 | 2,380932 | 2,380932 |

Setelah itu ditentukan tingkat kemiripan dokumen yang ditunjukkan pada tabel dibawah ini :

**Tabel 4. 6** Hasil Perhitungan *Cosine Similarity*

| No       | Q         | D1       | D2       |
|----------|-----------|----------|----------|
| Q D      | 7,5719252 | 5,668835 | 5,668835 |
| V        | 2,751713  | 2,380932 | 2,380932 |
| Q  V     | 4,564621  | 3,673841 | 3,673841 |
| Sumilary | 0,659405  | 0,56512  | 0,56512  |
| Nilai    | 65,94%    | 56,51%   | 56,51%   |

Berdasarkan hasil diatas maka Q12 yang dapat nilai tertinggi adalah dokumen D1.

### 4.3 Kelebihan dan Kekurangan Sistem

Sistem yang dibuat saat ini belum sepenuhnya sempurna, berdasarkan dari penelitian tentang Pengelompokan Publikasi Jurnal dan Laporan PKPM IIB Darmajaya, terdapat kelebihan dan kekurangan dalam sistem yang dibuat yaitu, sebagai berikut:

#### 4.3.1 Kelebihan Sistem

Adapun kelebihan sitem Pengelompokan Publikasi Jurnal dan Laporan PKPM IIB Darmajaya Menggunakan *Similarity Check* dengan Metode *Cosine* adalah sebagai berikut :

1. Sistem mudah digunakan sehingga mahasiswa IIB Darmajaya tidak lagi mengisi google form untuk mengumpulkan serta mengelompokan laporan mereka.
2. User dapat mengakses dimana saja yang terdapat internet. Sehingga dapat mempermudah user dalam melihat, serta mengupload laporan.
3. Tidak hanya melalui komputer atau laptop, user juga dapat membukanya melalui smartphone.

4. Terdapat juga mendownload data laporan yang sebelumnya jika user ingin melihat untuk menjadikan acuan dalam melakukan pembuatan laporan PKPM nya.

#### **4.3.2 Kekurangan Sistem**

Berikut kekurangan sitem Pengelompokan Publikasi Jurnal dan Laporan PKPM IIB Darmajaya Menggunakan *Similarity Check* dengan Metode *Cosine* adalah sebagai berikut :

1. Jika tidak ada akses internet User tidak dapat membuka website yang telah dibuat.
2. Admin yang menginput data laporan harus dilakukan satu persatu, sehingga merepotkan admin untuk memasukan data laporan.
3. Jika data training terlalu sedikit, Cosine Similarity tidak bisa bekerja dengan baik, dikarenakan data training yg menentukan data olahan sebelumnya. Agar hasil dari metode yg digunakan lebih efektif.