

PETUNJUK PENGGUNAAN

NLP PREPROCESSING TOOLKIT
– TEXT PREPROCESSING –

PETUNJUK PENGGUNAAN

1. Masuk ke halaman dashboard

The screenshot shows a web-based application interface. At the top, there's a header with user information (Spaces, Wafha, Metopen_NLP, Running) and navigation links (App, Files, Community, Settings). Below the header, the title "NLP Preprocessing Toolkit by Kelompok 1" is displayed. A descriptive text below the title states: "Aplikasi ini melakukan preprocessing teks Bahasa Indonesia yang mencakup: case folding, tokenization, stopword removal, normalization dan stemming." A central feature is a file upload area with the instruction "Unggah file CSV atau Excel" and a "Drag and drop file here" button, which includes a note "Limit 200MB per file • CSV, XLSX". To the right of the upload area is a "Browse files" button.

Di bagian tengah halaman, tersedia fitur unggah file yang memungkinkan pengguna mengimpor dataset dalam format CSV atau Excel (XLSX). Pengguna dapat mengunggah file dengan cara drag and drop atau dengan menekan tombol Browse files. Aplikasi ini mendukung ukuran file hingga 200 MB per file.

2. Preview Data Asli

Preview Data Asli

	id	text
0	1	Saya kecewa dengan pelayanan administrasi yang lambat
1	2	Petugas membantu dengan baik dan jelas
2	3	Proses pembuatan KTP terlalu lama
3	4	Pelayanan online cukup mudah digunakan
4	5	Sistem sering error dan tidak bisa diakses

Fitur ini berfungsi untuk memastikan bahwa file berhasil diunggah dan data dapat dibaca dengan benar oleh sistem. Pengguna dapat melihat contoh isi data sebelum melanjutkan ke proses berikutnya, seperti preprocessing atau analisis teks.

3. Pilih Kolom Teks

The screenshot shows a dropdown menu for selecting text columns. The menu lists "id", "text", and "text" again, with the second "text" option highlighted by a red border. Below the menu is a button labeled "Mulai Preprocessing".

Pada bagian ini, pengguna memilih kolom teks yang akan diproses. Pilih kolom text dari daftar yang tersedia, kemudian klik tombol Mulai Preprocessing untuk memulai proses pengolahan data.

4. Hasil Preprocessing

Preprocessing selesai!

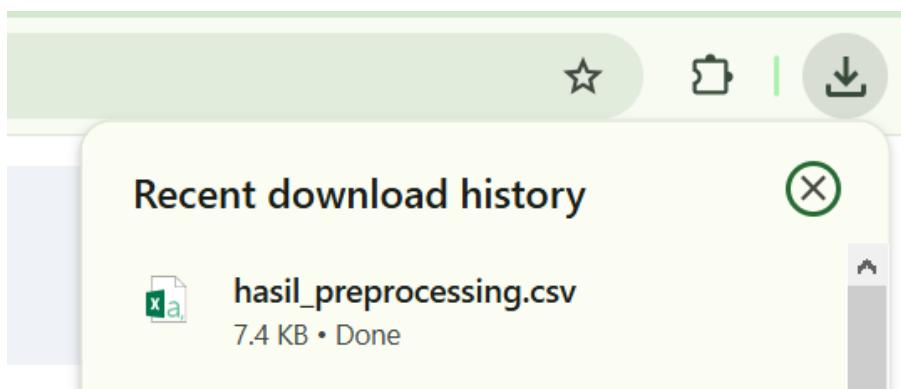
Hasil Preprocessing

	text	teks_bersih
0	Saya kecewa dengan pelayanan administrasi yang lambat	saya kecewa layan administrasi lambat
1	Petugas membantu dengan baik dan jelas	tugas bantu baik jelas
2	Proses pembuatan KTP terlalu lama	proses buat ktp terlalu lama
3	Pelayanan online cukup mudah digunakan	layan online cukup mudah guna
4	Sistem sering error dan tidak bisa diakses	sistem sering error tidak bisa akses

[Unduh Hasil Preprocessing \(CSV\)](#)

Setelah proses selesai, sistem menampilkan hasil preprocessing dalam bentuk tabel. Kolom teks_bersih menunjukkan teks yang telah melalui tahap pembersihan. Pengguna dapat mengunduh hasil preprocessing dengan menekan tombol Unduh Hasil Preprocessing (CSV).

5. Mengunduh Hasil Preprocessing



Setelah proses preprocessing selesai, sistem akan secara otomatis menghasilkan file keluaran dalam format CSV. File hasil tersebut dapat diunduh dan akan tersimpan di perangkat pengguna dengan nama hasil_preprocessing.csv, sebagaimana ditunjukkan pada gambar.

6. Melihat Hasil Preprocessing

A	B	C	D
id	text	teks_bersih	
1	Saya kecewa dengan pelayan saya kecewa layan administrasi lambat	saya kecewa layan administrasi lambat	
2	Petugas membantu dengan tugas bantu baik jelas	tugas bantu baik jelas	
3	Proses pembuatan KTP terlalu proses buat ktp terlalu lama	proses buat ktp terlalu lama	
4	Pelayanan online cukup mudah guna	layan online cukup mudah guna	
5	Sistem sering error dan tidak bisa akses	sistem sering error tidak bisa akses	
6	Pegawai kurang ramah saat layan kurang ramah saat layan	pegawai kurang ramah saat layan kurang ramah saat layan	
7	Informasi yang diberikan sar informasi beri sangat bantu	informasi yang diberikan sar informasi beri sangat bantu	
8	Antrian terlalu panjang dan t antri terlalu panjang tidak tertib	antrian terlalu panjang dan t antri terlalu panjang tidak tertib	
9	Pelayanan hari ini cukup me layan hari cukup muas	pelayanan hari ini cukup me layan hari cukup muas	
10	Pelayanan di kantor kelurahan kantor lurah sangat cepat ramah	pelayanan di kantor kelurahan kantor lurah sangat cepat ramah	

Setelah file hasil preprocessing diunduh dan dibuka, pengguna dapat melihat struktur data yang telah diproses dalam bentuk tabel. Dataset terdiri dari beberapa kolom, yaitu id, text, dan teks_bersih.

Kolom text berisi data teks asli sebelum diproses, sedangkan kolom teks_bersih berisi hasil preprocessing yang telah melalui tahapan seperti case folding, tokenisasi, dan pembersihan teks. Data pada kolom ini siap digunakan untuk proses analisis lanjutan atau pemodelan NLP.