**LAPORAN TUGAS CRAWLING DAN PREPROCESSING DATA TEXT**

**MATA KULIAH TEXT MINING & NATURAL LANGUAGE PROCESSING**

**Logo

Description automatically generated**

Tim Penyusun:

1. <NPM> <Nama Lengkap >
2. <NPM> <Nama Lengkap >
3. <NPM> <Nama Lengkap >
4. <NPM> <Nama Lengkap >
5. <NPM> <Nama Lengkap >

**PROGRAM STUDI SAINS DATA PROGRAM SARJANA**

**FAKULTAS SAINS & TEKNOLOGI**

**UNIVERSITAS TEKNOLOGI YOGYAKARTA**

**2025**

1. **Hastag yang digunakan:** #ramadhan
2. **Tuliskan langkah-langkah crawling data text:**
3. **Tuliskan/screenshotkan codingnya**

**Jawab:**

1. **Tuliskan/screenshotkan hasil crawlingnya**

**Jawab:**

****

1. **Berapa jumlah record yang didapatkan**

**Jawab: 6543 baris**

1. **Tuliskan langkah-langkah preprocessing data text:**
2. **Tuliskan/screenshotkan coding cleaning**

**Jawab:**

1. **Tuliskan/screenshotkan Case folding**

**Jawab:**

1. **Tuliskan/screenshotkan Tokenizing**

**Jawab:**

1. **Tuliskan/screenshotkan Filtering**

**Jawab:**

1. **Tuliskan/screenshotkan Stemming**

**Jawab:**

1. **Tuliskan/screenshotkan Simpan data text bersih**

**Jawab:**

1. **Lampirkan file RAW hasil Crawling dan file bersih hasil preprocessing**