*Full Stack Data Science Project*

|  |  |
| --- | --- |
| Name : | Risky Armansyah |
| NIM : | 09021282126055 |
| From : | Sriwijaya University |
|  |  |

Selamat pagi, siang, malam Risky, ini adalah dokumentasi yang akan kita gunakan untuk memantau dan mengamati perkembangan dari mega projek dalam rangka untuk memajukan diri maupun bangsa. Fokus utama kita adalah memanfaatkan data twitter yang biasanya digunakan masyarakat sebagai tempat untuk mengeluh. Berdasarkan data tersebut, kita harus melakukan analisis dan penyelesaianm masalah yang akan membantu bangsa. Karena merumuskan masalah masyarakat itu sulit, jadi fokus utama saat ini adalah melakukan dokumentasi apa yang kita lakukan tiap harinya terhadap perkembangan projek.

Berikut beberapa rumusan masalah yang ditemukan setelah brainstorming selama beberapa hari:

1. Sentiment analysis twit yang mengandung kata [allah, tuhan, islam, muhammad, dsb yang senandung dengan islam]. Sebagai fokus utama saat ini, yaitu scrapping dataset twitter, yang sudah dilakukan 2 cara pendekatan dan gagal (coba test lagi dengan docker). Model sentiment analysis yang akan digunakan mungkin dapat menggunakan model umum saja. Untuk tujuan selanjutnya, akan dibahas kembali.

Fokus utama terhadap masalah di atas / projek tersebut:

* 1. Cari cara untuk scraping twit twitter
  2. Terapkan sentiment analysis terhadap dataset twit tersebut, buat visualisasi yang mungkin berguna.

1. Mengenai permasalahan pangan yang dicetuskan, mungkin fokus utama saat ini adalah mencari dataset yang dapat menginterpretasikan keadaan pangan Indonesia. Kita dapat menggunakan berbagai teknik seperti scraping twitter, scraping web, data pemerintah, ataupun memberikan quisioner kepada masyarakat. Berdasarkan kumpulan beberapa data tersebut, kita dapat menganalisis keadaan pangan Indonesia, dan dapat mencari titik permasalahannya terletak di mana. Untuk penerapan penyelesaian permasalahan akan dilanjutkan nanti.
2. Masalah mindset rakyat Indonesia yang kekurangan asupan informasi sejak kecil. Notabenenya mindset yang kita punya sekarang itu merupakan cerminan dari apa yang kita alami. (?)

Terdapat 3 perumusan masalah, namun mari kita fokus ke permasalahan yang pertama, jadi fokus utama yaitu cari cara untuk scraping twit dari twitter (sumber informasi berharga).

Brainstorming terkait rumusan masalah, tujuan, dan manfaat dari analisis data twitter. Sampai saat ini, kita berhasil scrapping data tweet dari twitter menggunakan library twint python dengan total data sekitar 200rb tweet. Kita juga telah melakukan berbagai analisis serta preparasi data sesuai dengan kebutuhan dan hal yang terpikirkan saya saat ini. Akan tetapi, kita dihadapkan dengan permasalahan tujuan dan manfaat dari projek ini selain dari kepuasan akan pengetahuan sendiri. Jadi, mari kita fokus untuk mencari tujuan serta manfaat yang mungkin dapat diraih dari dataset tweet tersebut :

Rumusan masalaah :

* Adfs

Tujuan :

* Asdfa

Manfaat :