

The background features a complex arrangement of 3D geometric shapes, including rectangular prisms and cubes, in various shades of white, light gray, and orange. Some shapes have a fine grid texture. A network of orange dots connected by lines is visible in the upper center. A shield icon with a stylized 'S' is positioned on the right side. The overall aesthetic is modern and technological.

BINAR DSC

Challenge Report

Muhammad Agil Fahmi

Pendahuluan

Fenomena pengolahan data di Indonesia saat ini berkembang pesat. Cleansing data adalah salah satu tahap pengolahan data yang membutuhkan banyak waktu jika harus dilakukan secara manual.

Automasi pengolahan data diperlukan pada proses cleansing data untuk mempersingkat waktu pengolahan, namun diharapkan tetap mendapatkan hasil yang akurat.

Berdasarkan masalah di atas maka akan dilakukan rancang bangun API untuk melakukan cleansing data dengan menggunakan Bahasa pemrograman python dan SQLite sebagai database penyimpanan data (kamus) pembanding.

Metode Penelitian

Pada tahap awal penulis menentukan data yang akan diolah, Analisa ini akan dilakukan proses cleansing data hate speech yang bersumber dari Kaggle. Proses cleansing dilakukan dengan kondisi Ketika text input atau file input terdapat kata yg termasuk dalam kamus abusive atau kamus alay yang telah disimpan di database SQLite, maka secara otomatis kata tersebut akan disensor/diperbaiki susunan katanya. Semua proses cleansing akan dilakukan dengan bahasa pemrograman python berkolaborasi dengan Flask dan Swagger sebagai API. Pandas library dan Regex juga digunakan untuk membantu mempermudah proses cleansing data.

Hasil Dan Kesimpulan

Setelah dilakukan rancang bangun cleansing data API. Hasil setiap text yang diinput dan terdapat kata abusive yang match dengan kamus abusive maka kata tersebut akan disensor, jika dalam text terdapat kata alay yang match dengan kamus alay maka kata tersebut akan diperbaiki sesuai dengan EYD.

Proses cleansing dengan metode upload file masih gagal dilakukan dan diharapkan bisa dilakukan pada pengembangan API selanjutnya

The background of the entire image is an abstract, isometric 3D composition. It features numerous white and light gray rectangular blocks of varying heights and widths, some of which are topped with orange cubes. Thin, curved orange lines arc across the scene, connecting different points, suggesting a digital or networked environment. The overall color palette is dominated by white, light gray, and vibrant orange.

THANK YOU

