



Labelling Komentar pada YouTube

TUGAS BESAR MATA KULIAH PROYEK 1

PROGRAM STUDI D₄ TEKNIK INFORMATIKA
POLITEKNIK NEGERI BANDUNG

2020

IDENTITAS

| | |
|---------------|--|
| Nama Kelompok | : Candramawa |
| Manager | : Wendi Wirasta S.T., M.T. |
| Leader | : Kayis Shalahuddin (191524047) |
| Analyst | : Fachri Dhia Fauzan (191524041) Mohammad Rizky Maulidhan (191524049) |
| Designer | : Dzakira Fabillah (191524040) Nisa Shadrina (191524053) |
| Programmer | : Bima Putra Sudimulya (191524038) Kayis Shalahuddin (191524047) |

PENDAHULUAN

Latar belakang kami memilih judul ini disebabkan oleh banyaknya konten serta pengguna YouTube. Dilansir dari artikel milik IDN Times Jabar berjudul “*Data Statistik YouTube yang Jarang Orang Tahu*” pada tanggal 31 Agustus 2019, pengguna YouTube mencapai dua miliar per bulan dan terdapat lima ratus jam konten baru setiap menitnya. Oleh sebab itu, kami ingin mengetahui dari banyaknya konten serta pengguna YouTube ini, konten seperti apa yang menarik bagi pengguna dilihat dari komentar yang mereka berikan.

RENCANA TAHAPAN ANOTASI

SUMBER DATA

Sumber data yang kelompok kami gunakan dalam aplikasi ini berasal dari kolom komentar pada video YouTube dengan jumlah sampel data sebanyak 20 video dan setiap video 100 komentar. Data tersebut kami dapatkan melalui web scraping menggunakan Selenium.

JENIS ANOTASI

Pada tugas ini, kami akan menggunakan jenis anotasi categorization, yang menggunakan data berupa komentar pada YouTube dengan subdomain kata-kata sifat yang dapat menggambarkan perasaan pengguna YouTube ketika menonton konten tersebut.

TOOLS

Tools yang kami gunakan untuk anotasi data yaitu Label Studio.

RENCANA VISUALISASI DATA

STRUKTUR DATA

Visualisasi data hasil anotasi akan disimpan dengan format JSON, yang di-*dump* oleh sebuah dict python. Rencana isi list (Dapat berubah sesuai kebutuhan):

- isi komentar (tidak ditampilkan di aplikasi)
- Atensi komentar
- Judul video yang dikomentari
- link video
- Jumlah penonton

BAHASA PEMROGRAMAN

Bahasa pemrograman yang kami gunakan untuk visualisasi data hasil anotasi adalah python untuk pengumpulan data, pembuatan tampilan menggunakan PHP, CSS, JavaScript, dan HTML, pembuatan database menggunakan MySQL.

Library yang kami gunakan untuk visualisasi data hasil anotasi diantaranya, namun tidak terbatas pada :

- JSON

Sebagai media menyimpan data hasil anotasi data.

- Selenium

Sebagai media untuk melakukan web scraping.

ANTARMUKA DAN FUNGSIONALITAS

Rancangan antarmuka dari aplikasi visualisasi data yang kami buat :

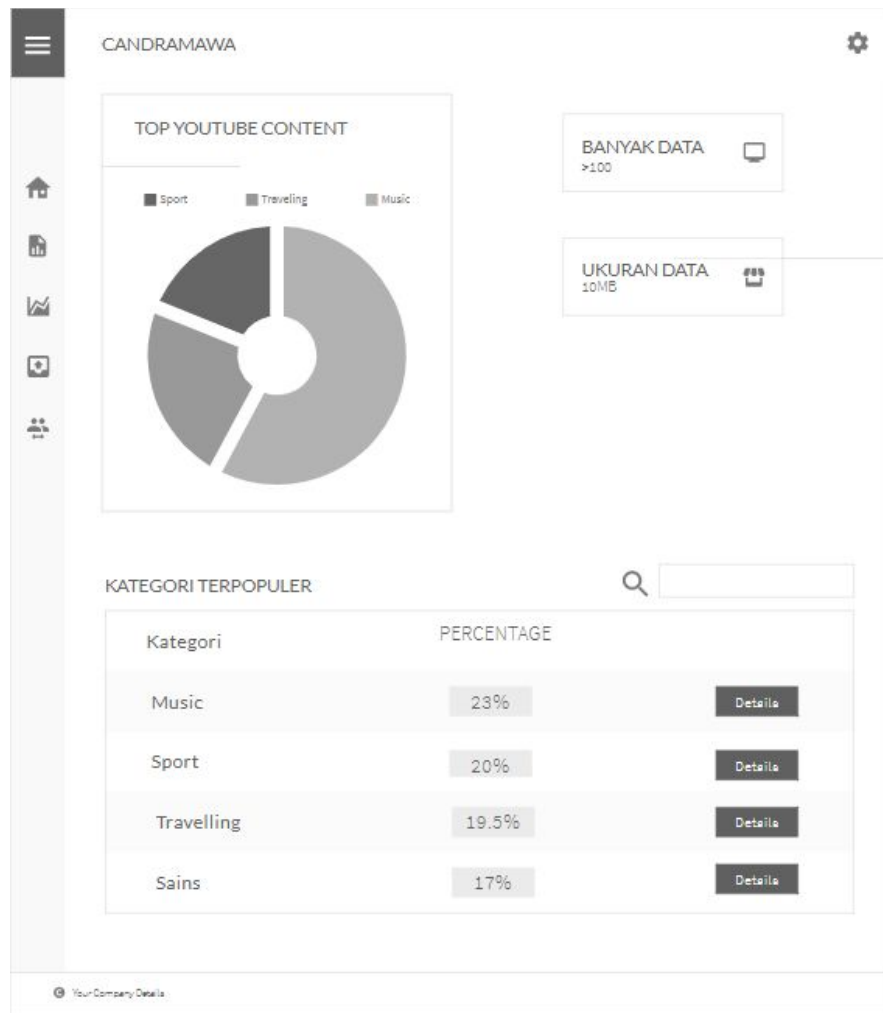


Figure 1 Halaman utama aplikasi

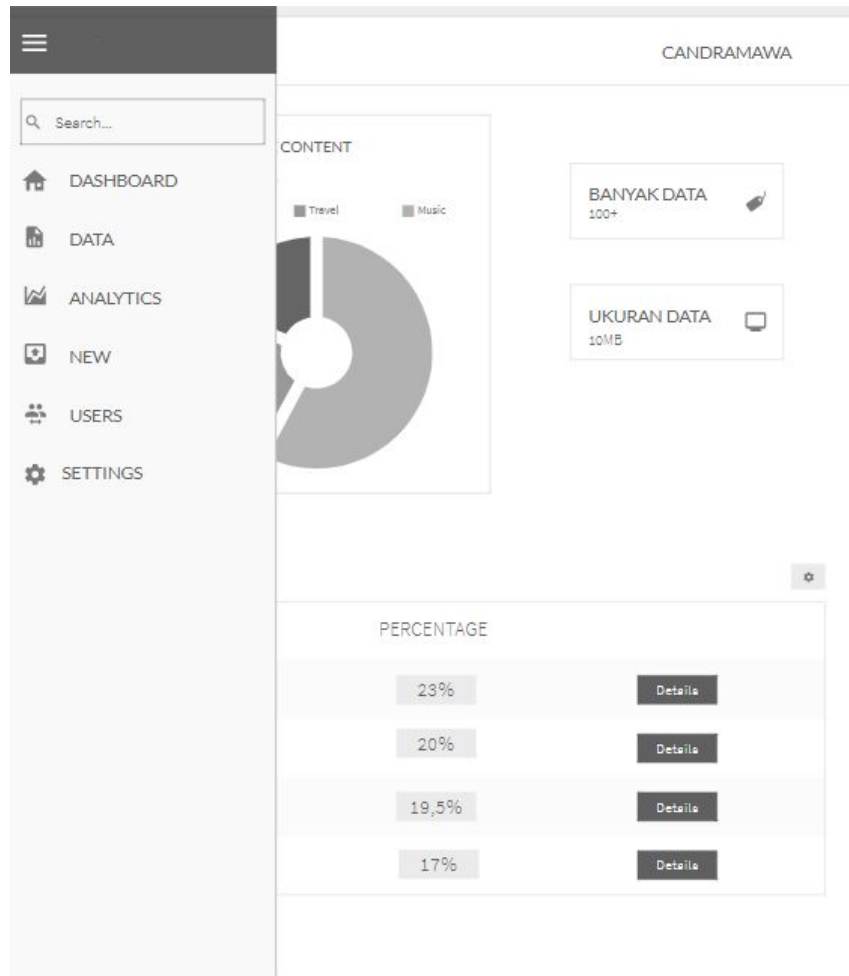


Figure 2 Halaman utama aplikasi +sidebar

Fitur-fitur yang disediakan pada aplikasi adalah

1. Menghitung banyak file yang digunakan
2. Menghitung total ukuran file
3. Menampilkan file terbaru yang dianotasi
4. Membuat grafik kategorisasi
5. Rekomendasi kategori konten.
6. Mencari kategori konten yang paling banyak/sedikit digemari pengguna YouTube.
7. Menampilkan jenis-jenis kategori konten.
8. Menampilkan jumlah peminat konten berdasarkan komentar dan jumlah penonton