

Anotasi Data dengan Segmentation pada Website Fantastic Film

Tugas besar mata kuliah Proyek 1

PROGRAM STUDI D4 TEKNIK INFORMATIKA

POLITEKNIK NEGERI BANDUNG

Bandung 2020

# **IDENTITAS**

Nama Kelompok : Fantastic Five

Manager : Trisna Gelar Abdillah

Leader : Laras Rasdiyani - 191524016

Analyst : Hanifah Ghina Nabila - 191524010

Designer : Farra Jessica - 191524008

Programmer : Ali Nurdin - 191524004

Linda Damayanti - 191524017

# **PENDAHULUAN**

Poster pada film mempunyai fungsi sebagai media promosi dari sebuah film dan dijadikan sebagai identitas film tertentu. Identitas dari film yang dimaksud seperti judul, sutradara, produser, pemeran, tanggal rilis, genre, dan lain sebagainya. Meskipun demikian, pada awalnya pemasangan dari sebuah poster film terbatas hanya di areal seputar bioskop saja, sehingga fungsi dari poster film sebagai media promosi terkadang tidak terealisasikan dengan sempurna.

Namun, di masa modern seperti sekarang, poster film dapat kita jumpai dimana saja. Dengan bantuan teknologi dan perkembangan sosial media pula, poster film dapat kita jumpai di berbagai platform seperti instagram, facebook, dan website khusus film.

Pada website online khusus film umumnya akan menyediakan berbagai informasi dan identitas terkait film tertentu. Seperti website IMDb, IMDb adalah singkatan dari *Internet Movie Database*, situs yang memiliki alamat web [www.imdb.com](http://www.imdb.com/) ini memuat berbagai informasi tentang film, program televisi dan video game, termasuk aktor, crew, dan karakter fiksi yang terdapat dalam tiga media hiburan visual tersebut. Saat ini Database IMDb telah diperluas untuk mencakup kategori tambahan, seperti demografis film tersebut, trivia, biografi, dan ringkasan cerita. Di dalam IMDb juga ada komunitas yang dapat berkontribusi langsung untuk menuangkan review tentang film dan memberikan rating pada film tersebut. Tidak hanya dari kaum awam, para pakar-pun juga mempunyai wadah sendiri untuk memberi rating dan menuangkan review secara profesional untuk film tersebut.

Dengan demikian penulis memilih untuk melakukan anotasi data berjenis *segmentation* dengan cara memberikan label pada segmen atau bagian kecil dari sebuah poster film yang didapat dari hasil scraping website IMDb. Pemilihan anotasi data ini bertujuan untuk mempermudah user dalam mendapatkan informasi mengenai dunia perfilman, sehingga dalam tugas besar mata kuliah proyek 1 ini penulis memilih judul “Anotasi Data dengan Segmentation pada Website Fantastic Film”.

# **RENCANA TAHAPAN ANOTASI**

1. SUMBER DATA

Sumber data yang kami ambil untuk tahap anotasi adalah melalui website IMDb, dengan penggunaan metode web scraping. Adapun data yang akan di scraping antara lain :

* Poster
* Judul
* Genre
* Tahun Rilis
* Sinopsis

1. JENIS ANOTASI

Jenis anotasi yang akan digunakan oleh pengembang aplikasi adalah *segmentation*. Pengembang aplikasi menggunakan anotasi data *segmentation* dengan cara memberi label pada setiap gambar berupa poster film. Label tersebut diambil berdasarkan genre dan ciri dari setiap film tersebut, dengan minimal sampel 50 file.

1. TOOLS

Dalam anotasi data dengan *segmentation* ini pengembang aplikasi menggunakan tools berupa Label Studio.

# **RENCANA VISUALISASI DATA**

1. STRUKTUR DATA

Anotasi yang dilakukan akan mengacu pada *image*, dalam hal ini berupa poster film yang kemudian akan diberi label perposternya berdasarkan ciri dari poster tersebut. Data hasil anotasi akan ditampilkan pada website Fantastic Film. Adapun data yang akan ditampilkan diantaranya :

* Poster (Hasil Sceraping website IMDb dan sebagai sumber anotasi)
* Judul (Hasil Scraping website IMDb)
* Genre (Hasil Scraping website IMDb)
* Ciri (Hasil anotasi data)
* Tahun Rilis (Hasil Scraping website IMDb)
* Sinopsis (Hasil Scraping website IMDb)

Dalam menentukan label atau segmen pada poster, pengembang menentukan berdasarkan :

* Benda
* Situasi
* Warna
* Manusia
* Perilaku objek

Adapun contoh data yang akan dianotasi misalnya poster pada film seperti berikut :



Pada film tersebut akan diberikan label seperti api, kegelapan, hitam, laki-laki, menatap, dan lain-lainnya.

Data hasil anotasi berupa labelling tersebut akan ditampilkan pada website Fantastic Film dengan nama “ciri”. Ciri ini kemudian akan ditampilkan lewat tabel bersama data - data lain hasil dari scrapping website IMDb, seperti yang diilustrasikan pada *Gambar 5. Tampilan halaman annotation.* Berikut struktur data yang dibutuhkan :

record film <

number: integer

judul: array of char

genre: array of genres

ciri : array of char

Tahun: integer

Sinopsis: array of char

>

record genres<

drama : array of char

romance :array of char

action :array of char

adventure :array of char

animation :array of char

biography :array of char

comedy :array of char

music :array of char

horror :array of char

sci-fi:array of char

documentary : array of char

fantasy : array of char

thriller : array of char

>

Struktur data diatas dapat ditambah, dikurangi, atau diganti sesuai dengan kebutuhan aplikasi nantinya.

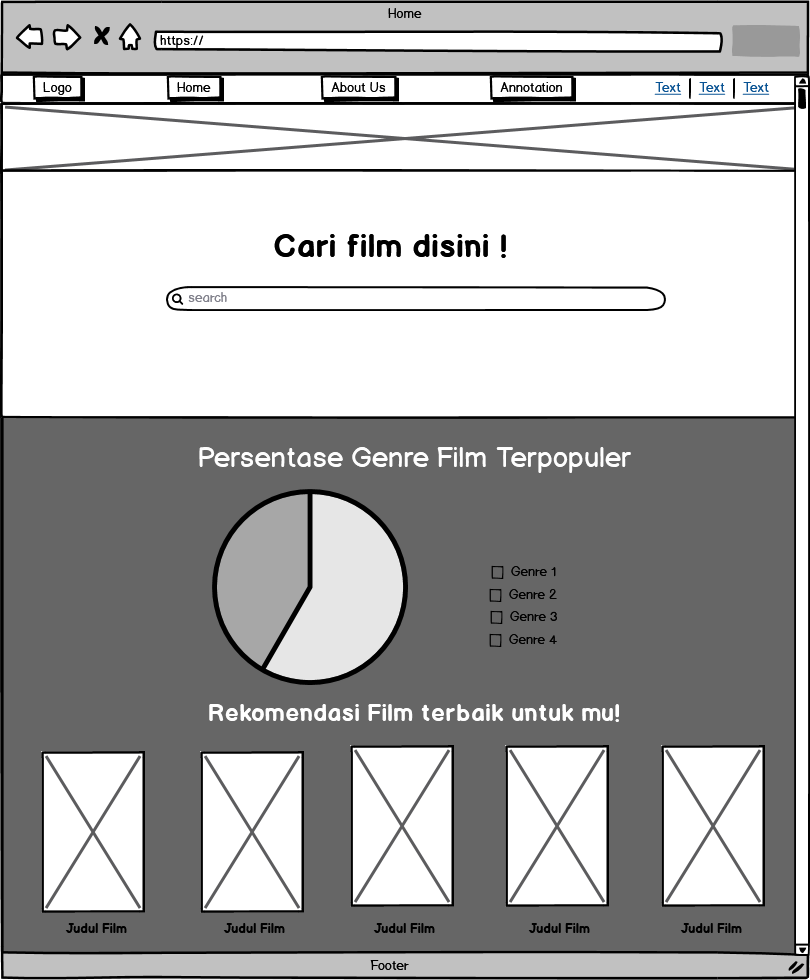
1. BAHASA PEMROGRAMAN

Bahasa pemrograman yang digunakan dalam pengerjaan proyek ini adalah Python. Dengan penggunaan metode web scraping dan penganotasian data *segmentation*, maka library yang akan kami pakai adalah:

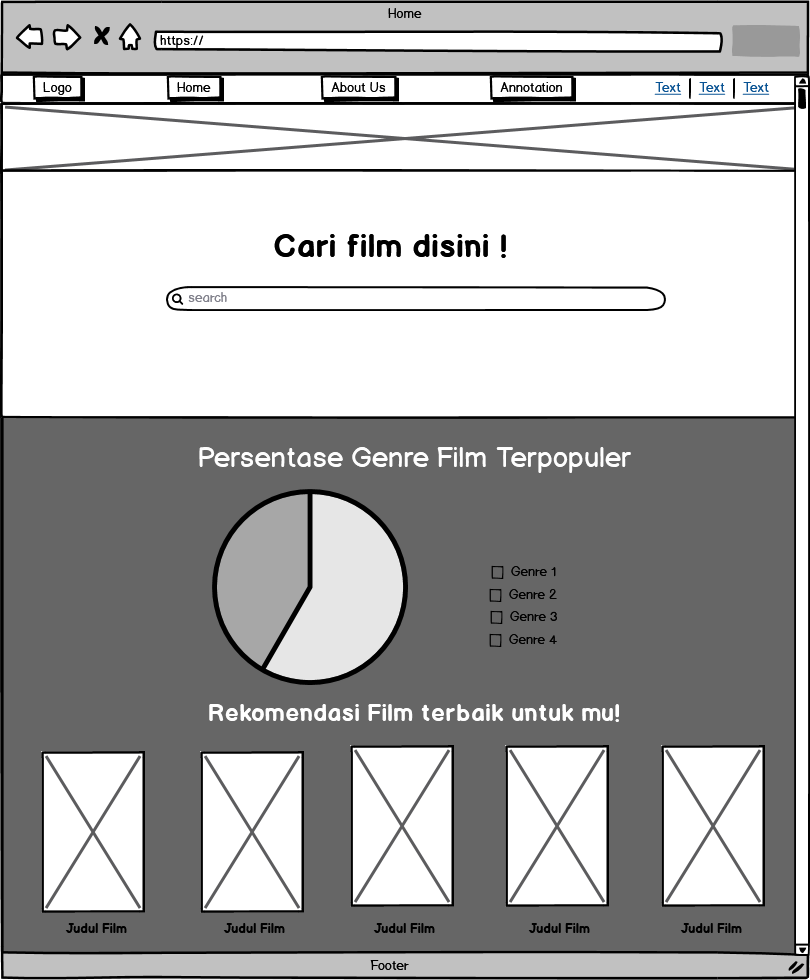
* Selenium
* Json
* Urllib
* Request
* BeautifulSoup4
* Scrapy
* Remote WebDriver

Library diatas dapat ditambah atau dikurangi sesuai dengan kebutuhan aplikasi nantinya.

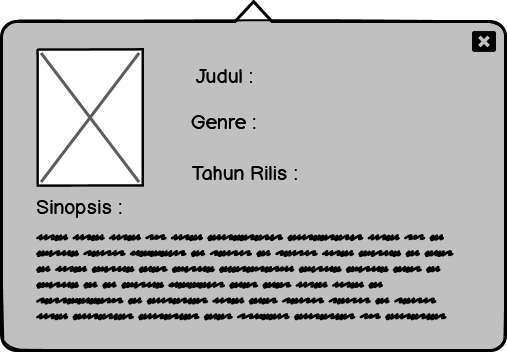
1. ANTARMUKA DAN FUNGSIONALITAS

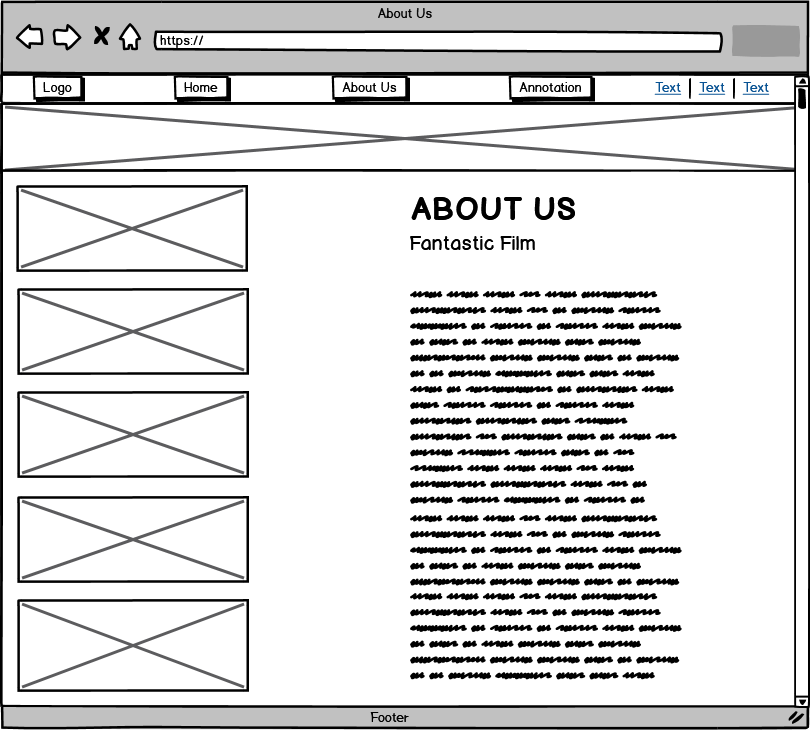


*Gambar 1. Tampilan halaman home sebelum input keyword pada search bar(jika tidak di scroll)*

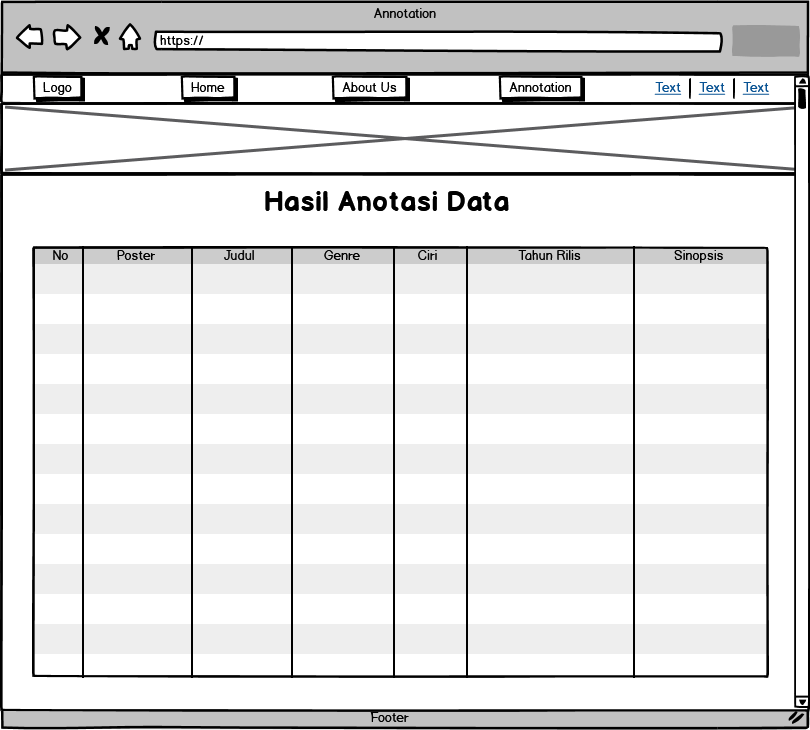
**

*Gambar 2. Tampilan halaman home sebelum input keyword pada search bar jika di scroll*

*Gambar 3. Tampilan pop up ketika poster film di klik*

**

*Gambar 4. Tampilan halaman about us*

**

*Gambar 5. Tampilan halaman annotation*

Fitur-fitur yang disediakan pada aplikasi adalah :

1. Memberikan informasi berupa diagram yang berisi persentase genre film terpopuler
2. Memberikan rekomendasi 5 film terbaik menurut website
3. Menampilkan data hasil anotasi berupa labelling pada setiap poster
4. Menampilkan data hasil scraping website berupa identitas film (Judul, Genre, Tahun Rilis, dan Sinopsis)
5. Mencari film berdasarkan labelling