

PROGRAM STUDI SARJANA SISTEM INFORMASI
PROPOSAL PROYEK AKHIR MATA KULIAH VISUALISASI DATA



MOVIE VISUALIZATION
INSTITUT TEKNOLOGI DEL

OLEH:

12S17003 Lavenia Situmorang
12S17009 Prince Ephraim Prabowo Silaban
12S17055 Josua Koko Sarasi Manurung
12S17062 Venny Handayani Sormin

INSTITUT TEKNOLOGI DEL
NOVEMBER 2020

DAFTAR ISI

DAFTAR ISI.....	ii
DAFTAR TABEL	iii
DAFTAR GAMBAR.....	iv
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Pengerjaan Proyek	1
1.3 Manfaat Pengerjaan Proyek	2
1.4 Ruang Lingkup.....	2
BAB 2. ISI.....	3
2.1 Pengumpulan Data	3
2.2 Data Preprocessing.....	4
2.2.1 <i>Data Cleaning</i>	4
2.2.2 <i>Case Folding</i>	4
2.2.3 Tokenizing	4
2.2.4 <i>Filtering</i>	4
2.2.5 <i>Stemming</i>	4
2.3 Menyusun Pertanyaan Visualisasi	5
2.4 What-Why-How Analysis Framework	5
BAB 3. RENCANA KERJA.....	6
DAFTAR PUSTAKA.....	7

DAFTAR TABEL

Tabel 1. Rencana Pengerjaan Proyek.....	6
---	---

DAFTAR GAMBAR

Gambar 1. Tahapan Pengerjaan Proyek	3
---	---

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Film adalah hasil karya seni berupa rangkaian gambaran hidup yang diputar sehingga menjadi sebuah ilusi gambar yang bergerak dan biasanya disajikan sebagai wujud suatu hiburan bagi manusia. Hasil karya seni berupa rangkaian gambar yang memiliki alur cerita tersebut menghasilkan gerakan berupa gerak-gerik layaknya kehidupan nyata dalam bentuk tanyangan video. Film juga sering disebut dengan *movie*. Saat ini, film (*movie*) merupakan salah satu bentuk seni yang sudah menjadi industri yang populer dimana hasil karya seni ini dapat digunakan untuk kepentingan bisnis dan hiburan. Terdapat berbagai jenis film (*movie*) dan salah satunya adalah anime. Anime adalah salah satu jenis film kartun atau animasi yang diproduksi oleh negara Jepang yang disukai oleh banyak orang. Banyaknya jenis (*genre*) anime yang diterbitkan menjadi salah satu hal yang menarik bagi para penggemar atau pecinta anime. Jenis (*genre*) anime tentunya memiliki nilai (*rating*) yang berbeda-beda dari setiap orang yang pernah melihatnya. Sebagai salah satu penggemar anime, menemukan informasi mengenai anime yang terbaik atau favorit adalah hal yang menyenangkan. Hal ini juga menjadi salah satu hal yang penting bagi pelaku bisnis dimana informasi yang didapat memungkinkan pelaku bisnis menjadikannya sebagai informasi untuk pengembangan bisnisnya. Penulis berasumsi bahwa respon dari strategi pemasaran dapat diperkirakan menggunakan jumlah *review* dari ulasan sebuah film (*movie*). Semakin banyak ulasan berarti semakin intensif dan respon positif terhadap pemasaran sebuah film (*movie*). Untuk hal itu kami mengusulkan pembuatan visualisasi.

Visualisasi data merupakan bagian yang dapat diterapkan untuk tujuan tersebut. Dengan adanya visualisasi data diharapkan dapat mengkomunikasikan informasi secara jelas dan efisien melalui informasi grafik yang dipilih.

1.2 Tujuan Pengerjaan Proyek

Adapun yang menjadi tujuan dari pengerjaan proyek ini adalah sebagai berikut:

1. Menampilkan nama anime yang tergolong *genre*:
 - Action, Fantasy, Game.
 - Drama, Josey, Romance.
 - Comedy, Hentai, Magic.
 - Mecha, Mystery

2. Menampilkan anime yang memiliki type *music* dan *special*.
3. Menampilkan 10 anime terfavorit berdasarkan *members* yang paling banyak.
4. Menampilkan 10 anime terbaik berdasarkan rating tertinggi.

1.3 Manfaat Pengerjaan Proyek

Manfaat dari proyek ini adalah menampilkan beberapa informasi data mengenai anime yang disajikan dalam suatu visualisasi data berbentuk *dashboard*.

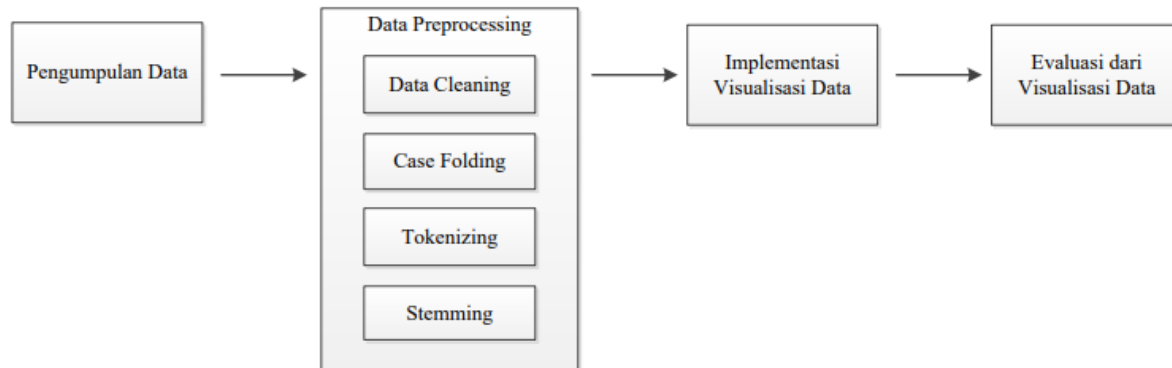
1.4 Ruang Lingkup

Batasan dalam penelitian ini adalah sebagai berikut:

1. Sumber data berasal dari Kaggle. Data terdiri dari 12.294 baris data. Data anime terdiri dari tujuh atribut yaitu *anime_id*, *name*, *genre*, *type*, *episode*, *rating* dan *members*.
2. Data yang digunakan adalah dataset anime dan berbentuk *csv*.
3. Data yang divisualisasikan akan disajikan dalam bentuk grafik dengan menggunakan Tableau 2020.

BAB 2. ISI

Arsitektur sistem pengerjaan proyek visualisasi data movie visualization yang akan dilakukan dapat dilihat pada gambar dibawah ini.



Gambar 1. Tahapan Pengerjaan Proyek

2.1 Pengumpulan Data

Dataset yang digunakan pada proyek ini merupakan *dataset movie* anime yang terdiri dari 12294 baris data yang diambil dari kaggle.com dan telah diberikan label untuk setiap genre dari movie tersebut, apakah termasuk kategori *action*, *comedy*, *romance*, dan *adventure*. *Dataset* tersebut merupakan data yang bersifat statis karena dalam format file CSV (*Comma Separated Values*).

(1) Anime_ID

Anime_ID merupakan ID dari setiap anime (*Movie*) yang merupakan nilai unik.

(2) Name

Name merupakan judul dari anime tersebut

(3) Genre

Genre adalah aliran dari anime tersebut.

(4) Episodes

Episodes merupakan banyaknya jumlah episode dari movie tersebut

(5) Rating

Rating merupakan penilaian yang terstandarisasi dari movie tersebut

(6) Members

Members merupakan banyaknya penonton movie tersebut.

2.2 Data Preprocessing

Dataset yang didapat untuk penelitian pada dasarnya tidak langsung siap digunakan. Hal ini biasa terjadi karena kesalahan transmisi dan kesalahan yang dilakukan manusia (*human error*). Pra-pemrosesan dilakukan dengan tujuan untuk mendapatkan informasi yang dibutuhkan dari dataset yang tersedia. Teknik pra-pemrosesan yang akan diterapkan pada dataset akan disesuaikan dengan kebutuhan. Adapun teknik yang diterapkan adalah sebagai berikut:

2.2.1 Data Cleaning

Data cleaning diperlukan dalam proses *data mining*. Tahap awal dari pra-pemrosesan data adalah melakukan *data cleaning*. *Data cleaning* adalah proses menyiapkan data untuk dilakukan analisis dengan cara menghapus atau memodifikasi data yang salah, tidak relevan, duplikat dan tidak terformat. Penerapan *data cleaning* akan meliputi pengisian data yang kosong (*missing value*), mengidentifikasi atau menghilangkan *outlier* dan menghilangkan inkonsistensi pada data.

2.2.2 Case Folding

Case folding adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf ‘a’ sampai dengan ‘z’ yang diterima. Karakter selain huruf dihilangkan dan dianggap delimiter.

2.2.3 Tokenizing

Pada tahap ini dilakukan proses pemisahan teks menjadi potongan-potongan kata yang ada pada dataset. Proses tersebut dilakukan dengan memisahkan teks yang ada di datasets menjadi potongan-potongan kata ataupun tanda baca dan titik.

2.2.4 Filtering

Pada tahap ini dilakukan pengambilan kata-kata yang penting dari hasil tokenisasi. Bisa menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). *Stoplist* / *Stopword* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*. Contoh *stopwords* adalah “yang”, “dan”, “di”, “dari” dan seterusnya.

2.2.5 Stemming

Pada tahap ini dilakukan pencarian *root* kata dari tiap kata hasil *filtering*. Setelah mendapatkan hasil pencarian *root* kata maka akan dilakukan pengembalian berbagai bentukan kata kedalam suatu representasi yang sama.

2.3 Menyusun Pertanyaan Visualisasi

Setelah analisis data dilakukan selanjutnya proses implementasi. Implementasi visualisasi data dilakukan dengan menggunakan menggunakan tools Tableau atau Gephi. Dengan menerapkan visualisasi data, penulis akan menghasilkan sebuah *dashboard* yang informatif sehingga dapat mempermudah seseorang untuk mendapatkan informasi yang dibutuhkan.

2.4 What-Why-How Analysis Framework

Pada tahap ini akan dilakukan penilaian akan hasil visualisasi data yang telah dihasilkan selama pengerjaan proyek. Tipe evaluasi untuk mengevaluasi hasil dari visualisasi data akan disesuaikan dengan proses pengerjaan proyek. Adapun tipe evaluasi yang dipakai yaitu tipe kuantitatif: *formative - usability studies* dengan mekanisme.

BAB 3. RENCANA KERJA

Adapun rencana pengerjaan proyek yang akan dilakukan dapat dilihat pada tabel dibawah ini.

Tabel 1. Rencana Pengerjaan Proyek

Kegiatan	PIC	Bulan								
		November			Desember				Januari	
		W2	W3	W4	W1	W2	W3	W4	W1	W2
Menentukan Ide dan judul	Venny, Prince, Lavenia, Joshua									
Mengumpulkan Data	Venny, Prince, Lavenia, Joshua									
Pengerjaan proposal bab 1	Venny, Lavenia									
Pengerjaan proposal bab 2	Joshua, prince									
Pengerjaan proposal bab 3	Lavenia, Venny									
Pengerjaan laporan akhir analisis	Venny, Prince, Lavenia, Joshua									
Pengerjaan laporan akhir desain	Venny, Prince, Lavenia, Joshua									
Pengerjaan laporan akhir implementasi	Venny, Prince, Lavenia, Joshua									
Pengerjaan laporan akhir evaluasi	Venny, Prince, Lavenia, Joshua									
Pengumpulan proyek	Venny, Prince, Lavenia, Joshua									

DAFTAR PUSTAKA

- [1] Chen, T., Lu, A., & Hu, S. M. (2012). Visual storylines: Semantic visualization of movie sequence. *Computers & Graphics*, 36(4), 241-249.
- [2] Lin, D. S., Kan, A., Gao, J., Crampin, E. J., Hodgkin, P. D., & Naik, S. H. (2018). DiSNE movie visualization and assessment of clonal kinetics reveal multiple trajectories of dendritic cell development. *Cell reports*, 22(10), 2557-2566.