

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pantun merupakan salah satu bentuk puisi yang populer di berbagai budaya di seluruh dunia. Dalam bentuknya yang sederhana namun berirama, pantun dapat mengungkapkan ide, perasaan, dan cerita dengan cara yang kreatif dan berkesan. Dahulu, pantun digunakan untuk melengkapi percakapan sehari-hari. Bahkan saat ini masih digunakan oleh sebagian besar masyarakat Melayu di pedesaan. Pantun digunakan dalam pidato oleh tokoh adat dan tokoh masyarakat, oleh pedagang yang menjual produknya, oleh orang yang mengalami musibah, dan oleh orang yang ingin mengungkapkan kebahagiaan.

Dalam konteks ini, pantun populer sering kali menjadi *viral* di media sosial, menciptakan tren dan menginspirasi pengguna lainnya untuk berpartisipasi dalam kegiatan berpantun. Pantun-pantun tersebut dapat berasal dari berbagai topik, mulai dari kisah cinta, humor, peristiwa aktual, hingga kritik sosial. Melalui pantun, pengguna dapat mengekspresikan diri mereka dengan gaya yang unik dan menciptakan koneksi emosional dengan *audiens* mereka.

Di era digital saat ini, dengan berkembangnya media sosial, pantun telah menemukan tempatnya sebagai bentuk ekspresi yang populer di *platform*- seperti *twitter*, *instagram*, *facebook* atau *platform* lainnya. Media sosial memberikan wadah yang luas bagi pengguna untuk berbagi pantun-pantun mereka, memperoleh apresiasi dari teman-teman dan pengikut mereka, serta terlibat dalam interaksi berpantun yang menarik.

Kesulitan dalam melacak dan mengidentifikasi secara manual pantun-pantun populer di media sosial menjadi permasalahan yang umum terjadi di kalangan individu. Hal ini dikarenakan pantun yang di *posting* setiap harinya di berbagai platform media sosial berada dalam skala besar. Oleh karena itu,

mencari pantun-pantun populer diantara ribuan atau bahkan jutaan *postingan* menjadi tugas yang sangat rumit dan memakan waktu. Selain itu, media sosial juga memiliki algoritma yang kompleks untuk menampilkan konten kepada pengguna. Akibatnya, pantun-pantun populer mungkin tidak muncul dengan mudah di *feed* pengguna atau sulit untuk ditemukan melalui pencarian.

Proses manual dalam mencari pantun-pantun populer memakan waktu yang signifikan dan tidak efisien. Pengguna harus melakukan pencarian, membuat berbagai *postingan*, dan menganalisis konten secara manual. Hal ini dapat menjadi tugas yang melelahkan dan membatasi kreativitas serta produktivitas individu yang ingin menciptakan pantun baru.

Berdasarkan hasil *survey* melalui wawancara dan keluhan di sosial media terdapat sebanyak 75% responden mengungkapkan bahwa mereka memiliki keterbatasan dalam pemahaman tentang struktur pantun. Para pengguna tidak terbiasa dengan pola rima dan irama yang terdapat dalam pantun. Bahkan beberapa responden mengaku tidak tahu bahwa pantun terdiri dari empat baris dengan pola rima tertentu. Sehingga kurangnya pengetahuan tentang struktur pantun ini dapat menghambat kemampuan mereka dalam menciptakan pantun yang harmonis dan sesuai dengan aturan pantun tradisional.

Tidak hanya itu, sebanyak 80% responden juga mengaku menghadapi kesulitan dalam memahami gaya atau kriteria pantun yang baik dan cara menghasilkan ide-ide kreatif untuk pantun mereka. Beberapa responden juga mengungkapkan ke tidak pahaman tentang penggunaan kata-kata khas dalam pantun, seperti penggunaan peribahasa, kemudian kesulitan dalam menghubungkan gagasan atau membuat metafora yang menarik dalam pantun. Tantangan ini membuat para responden merasa terhambat dalam menciptakan pantun yang *orisinil* dan unik.

Twitter merupakan media sosial jenis *microblogging* (blog kecil) yang didirikan oleh *Jack Dorsey* (2006). Yang unik dari Twitter adalah tweet atau pesan Twitter memiliki panjang maksimal 140 karakter. Perayapan data di Twitter

adalah proses mengambil atau mengunduh data dari server Twitter menggunakan Twitter API berupa data pengguna dan data *tweet*. Aplikasi ini menggunakan dan mengumpulkan pantun-pantun yang sedang *viral*, banyak dibagikan, atau menarik perhatian pengguna.

Setelah data terkumpul, analisis pantun populer akan dilakukan untuk memahami pola-pola atau karakteristik yang membuat pantun tersebut diminati oleh pengguna media sosial. Analisis ini dapat mencakup analisis teks, pengenalan pola, atau metode pemodelan data untuk mengidentifikasi kata-kata yang sering digunakan, gaya penulisan yang menarik, atau tema-tema yang populer dalam pantun-pantun tersebut.

Data *mining* adalah teknik analisis data yang digunakan untuk mengekstrak pengetahuan yang tersembunyi atau pola-pola yang mungkin sulit ditemukan dengan cara-cara konvensional, Implementasi Data Mining pada *Tweet* Pantun merupakan proses penggunaan teknik data *mining* untuk menggali informasi yang berharga atau pola-pola yang menarik dari kumpulan *tweet* yang berisi pantun.

Klasifikasinya dengan cara memproses pengelompokan objek atau data ke dalam kategori atau kelas yang berbeda berdasarkan karakteristik atau atribut tertentu. Tujuan utama dari klasifikasi adalah untuk mengidentifikasi atau memprediksi kelas atau kategori dari objek atau data baru berdasarkan pengalaman yang telah ada pada data yang telah diklasifikasikan sebelumnya.

Berdasarkan penelitian yang telah ada sebelumnya, penelitian ini mencoba melakukan analisis model klasifikasi yang dapat mengidentifikasi dan mengategorikan *tweet* yang mengandung pantun di Twitter. Data tersebut akan diproses dengan data *mining* untuk lebih mudah mengklasifikasikan dalam kategori pantun dan bukan pantun.

Dalam penelitiannya, Dianati Duei Putri dkk (2022) melakukan *Crawling Twitter* menggunakan *query* untuk mengambil *tweet* dari pengguna Twitter.

Proses Crawling menggunakan *library tweepy Python* dan *Twitter API*. Dengan demikian, hasil penelitian menunjukkan skor akurasi sebesar 0,8 yang berarti sistem dapat memprediksi 20% dari total data pengujian dengan akurasi 80%. Analisis sistem kemudian menurunkan *rating DPR tweet Twitter* menjadi 95 positif, 693 netral, dan 758 negatif dari 1.546 hasil indeks data. Dengan demikian, dapat disimpulkan bahwa sistem penganalisis sentimen berhasil dirancang untuk menerima *tweet* dari Twitter tentang aktivitas DRP.

Sepyan Purnama Kristanto, dkk (2021) dalam penelitiannya menunjukkan bahwa metode pengumpulan datanya menggunakan model web *mining* yang memiliki salah satu metode terbaik dalam mengekstraksi data dari web yaitu *deep-first*, dimana data diekstraksi berdasarkan hubungan antar data untuk mendapatkan data yang akurat. Hasil penelitian menunjukkan tingkat keberhasilan 97% untuk alamat, 99% untuk kota, 70% untuk *rating*, dan 60% untuk ulasan di 211 data set untuk mengekstrak informasi dari empat variabel *crawling*.

## **1.2 Rumusan Masalah**

Pengguna kesulitan untuk membuat atau mengidentifikasi pantun-pantun secara manual dikarenakan jumlah pantun yang diunggah setiap hari sangat besar serta tidak adanya alat yang efektif yang mendukung proses pencarian pantun sehingga dibutuhkan Ai untuk membantu menganalisis dan mengklasifikasi tweet yang mengandung pantun di twitter.

## **1.3 Batasan dan Asumsi**

Berdasarkan identifikasi masalah diatas maka permasalahan dibatasi agar penelitian ini terarah pada Pengembangan Data di ambil dari *twitter* menggunakan teknik *crawling* untuk Analisis Pantun Populer di Media Sosial. Dengan adanya permintaan dan kebutuhan yang cukup dari pengguna untuk

memiliki alat otomatis yang dapat membantu mereka dalam pembuatan pantun, asumsi ini didasarkan pada fakta bahwa metode ini dapat membantu menghemat waktu dan meningkatkan produktivitas dalam menciptakan pantun, serta memudahkan pengguna untuk mengikuti tren dan terhubung dengan komunitas berpantun di media sosial.

#### **1.4 Tujuan Penelitian**

Adapun tujuan penelitian ini dilakukan untuk mengembangkan model klasifikasi yang dapat mengidentifikasi atau memprediksi kelas dan mengategorikan *tweet-tweet* yang mengandung pantun di Twitter. dengan menggunakan model *Support Vector Machine* (SVM) diharapkan dapat memisahkan *tweet-tweet* tersebut ke dalam tiga kategori yang telah ditentukan: "pantun," "bukan pantun," dan "pantun gambar."

#### **1.5 Manfaat Penelitian**

Penelitian ini memungkinkan penulis untuk memudahkan dan memberikan manfaat dalam pengembangan teknik data *mining* untuk memprediksi suatu kelas atau label berupa teks di media *social* . Hasil penelitian dapat digunakan sebagai dasar untuk penelitian lebih lanjut dalam metode data *mining* yang lebih canggih dengan menggunakan *platform Twitter*.

#### **1.6 Sistematika Penulisan**

Sistematika penulisan skripsi ini mempunyai struktur sebagai berikut:

##### **1. BAB I : PENDAHULUAN**

Meliputi latar belakang masalah, rumusan masalah, batasan dan asumsi tujuan penelitian, manfaat penelitian, metodologi penulisan.

##### **2. BAB II : TINJAUAN PUSTAKA DAN LANDASAN TEORI**

Memuat informasi tinjauan literatur penelitian terdahulu dan dasar teori perancangan sistem dari jurnal dan sumber terpercaya.

##### **3. BAB III : Metode Penelitian**

Memuat alat penelitian, bahan penelitian dan proses penelitian

#### 4. BAB IV : HASIL DAN PEMBAHASAN

Menjelaskan hasil penelitian, analisis dan pembahasan penelitian yang dilakukan.

#### 5. BAB V : KESIMPULAN DAN SARAN

Berisi kesimpulan tentang hasil umum penelitian yang dilakukan sesuai tujuan penelitian dan menyajikan saran penulis untuk penelitian selanjutnya.

#### 6. BIBLIOGRAFI

Berisi daftar sumber literatur dan teori yang penulis gunakan sebagai bahan diskusi dalam penyusunan skripsi ini.