

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada era digital saat ini hiburan di internet adalah hiburan yang paling banyak dikunjungi dikarenakan hiburan di internet mudah diakses dan memiliki biaya yang cukup murah. Salah satu hiburan yang meningkat pada era digital saat ini adalah hiburan membaca, hal ini dikarenakan saat akses internet masih sulit mencari hiburan dalam bentuk karya tulis masih sangat tergolong mahal untuk membeli buku tersebut apalagi jenis-jenis buku seperti buku pendidikan dan buku novel.

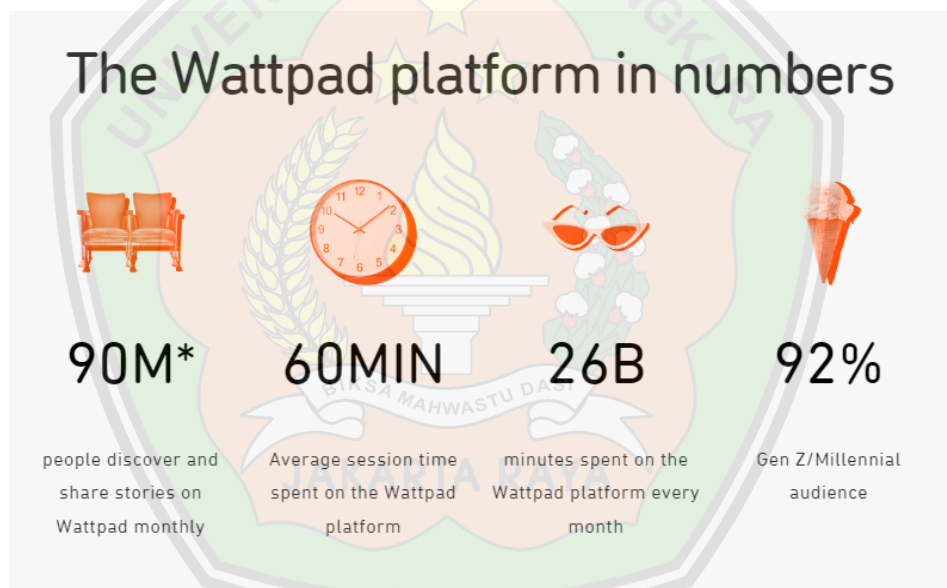
Saat ini banyak platform terkenal yang menyediakan hiburan dalam bentuk karya tulis, salah satu platform itu adalah Wattpad, sebuah platform cerita sosial yang dikenal luas, didirikan pada tahun 2006 oleh Allen Lau dan Ivan Yuen dengan tujuan membuat membaca dapat diakses melalui perangkat seluler. Pada tahun-tahun awalnya, Wattpad menarik pengguna dengan menawarkan buku-buku domain publik seperti *Pride and Prejudice*. Meskipun awalnya hanya menghasilkan pendapatan kecil, Wattpad berhasil mendapatkan popularitas dan pendanaan yang signifikan dalam waktu singkat [1].

Pada tahun 2011, Wattpad tersedia di berbagai perangkat seperti iPhone, Android, dan Blackberry, serta menarik satu juta pengguna terdaftar dan pendanaan lebih dari \$4,1 juta. Perkembangan pesat berlanjut, dan pada tahun 2014, Wattpad berhasil mengumpulkan USD\$17,3 juta dalam putaran pendanaan seri B. Pengguna Wattpad tumbuh dengan cepat, dan platform ini melampaui satu miliar menit dihabiskan per bulan. Buku berjudul "After" menjadi viral di Wattpad, yang kemudian diterbitkan oleh Simon & Schuster dan menjadi bestseller New York Times.

Pada tahun 2017, Wattpad mencatat unggahan ke-400 juta dan meluncurkan berbagai inisiatif baru seperti Wattpad Brand Partnerships, Wattpad WEBTOON Studios, dan program Wattpad Stars. Platform ini dianugerahi beberapa penghargaan inovasi dan pengusaha terbaik. Pada tahun 2019, Wattpad

memperkenalkan Paid Stories, yang memungkinkan penulis mendapatkan dukungan finansial langsung dari pembaca. Selain itu, Wattpad meluncurkan Wattpad Books dan ikut memproduksi film panjang pertama berdasarkan cerita viral dari platform mereka.

Saat ini, Wattpad adalah platform cerita sosial terbesar di dunia, dengan jutaan pengguna yang menghabiskan total 26 miliar menit di platform setiap bulan. Wattpad terus bekerja sama dengan industri hiburan untuk memproduksi cerita untuk berbagai media, serta memiliki kantor pusat di Toronto dan kantor satelit di seluruh dunia, termasuk di Halifax, Nova Scotia. Pernyataan ini dapat dilihat company.wattpad.com dan menampilkan data seperti berikut:



Gambar 1. 1 Pencapaian Wattpad [2]

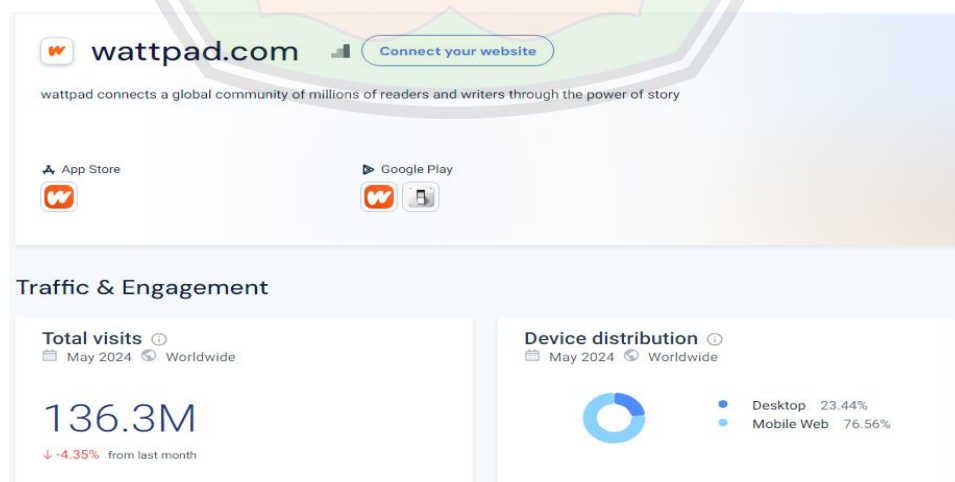
Wattpad merupakan platform cerita sosial yang memiliki pengguna aktif yang besar dan tingkat keterlibatan yang tinggi. Berikut adalah beberapa statistik kunci yang menggambarkan penggunaan dan demografi platform ini:

Tabel 1. 1 Pencapaian Wattpad

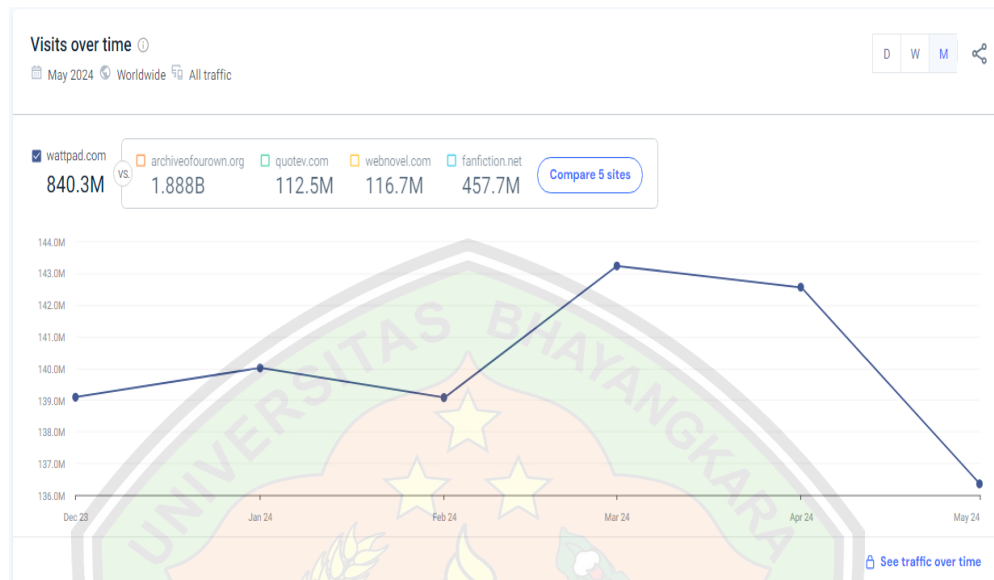
Statistik	Nilai	Keterangan
Pengguna Aktif Bulanan	90 juta	Jumlah pengguna yang aktif setiap bulan
Rata-rata Waktu Sesi	60 menit	Rata-rata durasi waktu yang dihabiskan oleh pengguna dalam satu sesi
Total Waktu Penggunaan Bulanan	26 miliar menit	Jumlah total menit yang dihabiskan oleh semua pengguna di platform setiap bulan
Demografi Pengguna	92%	Persentase pengguna yang berasal dari Generasi Z dan Milenial

Sumber : Company Wattpad [2]

Salah satu website analisis yang bernama pro.similarweb.com telah melakukan reset terhadap platform wattpad, berikut datanya :

**Gambar 1. 2** Data Pengunjung Bulain Mei 2024 [3]

Dari hasil reset yang dilakukan bahwa pada bulan mei tahun 2024 wattpad memiliki pengunjung sebanyak 136,3 juta pengunjung dengan pengguna desktop 23,44% dan pengguna mobile web 76,56%.



Gambar 1. 3 Grafik Pengunjung Bulan Desember 2023 – Mei 2024 [3]

Pada bulan desember 2023 sampai mei 2024 platform wattpad ini sudah mendapatkan total pengunjung sebanyak 840,3 juta pengunjung. Berdasarkan data tersebut platform wattpad memiliki banyak penulis yang membuat karya di platform wattpad tersebut. Membuat karya tulis di wattpad dapat dilakukan oleh siapapun dikarenakan membuat karya tulis di wattpad itu gratis dan hanya memerlukan email saja untuk mendaftar, dengan banyaknya penulis pada wattpad membuat algoritma pencarian di wattpad hanya menampilkan karya tulis yang terkenal dan karya tulis terkenal ini mudah untuk ditemukan hanya dengan memasuki daftar genre saja, sedangkan karya tulis yang kurang terkenal tidak memiliki genre dan hanya diberikan *symbol hashtag* hal ini membuat sulit bersaing untuk sebagian penulis yang belum ditentukan genrenya.

Berdasarkan masalah tersebut platform wattpad memiliki potensi untuk dijadikan tempat pengambilan sampel data novel berbahasa Indonesia. Data novel tersebut akan diolah untuk menentukan seberapa banyak persentase kalimat dalam novel tersebut berdasarkan kalimat-kalimat yang telah disesuaikan oleh data genre

tertentu. Proses ini akan menggunakan *Natural Language Processing* (NLP) sebagai tahapan pre-processing terhadap data novel agar bisa digunakan untuk klasifikasi dengan data genre yang akan dilakukan uji dan latih menggunakan algoritma *Support Vector Machines* (SVM). Pemilihan NLP pada penelitian ini dikarenakan NLP adalah cabang dari kecerdasan buatan (AI) yang fokus terhadap interaksi manusia melalui bahasa, dan juga NLP ini memiliki kemampuan multibahasa sehingga memungkinkan penelitian ini dapat melakukan klasifikasi terhadap novel-novel yang menggunakan bahasa Indonesia [4]. Penggunaan algoritma SVM pada penelitian ini dikarenakan SVM memiliki fleksibilitas dibandingkan algoritma lain dalam berbagai jenis data, terutama terhadap data novel yang biasanya isi novel memiliki isi data dari kecil, sedang, dan banyak tergantung jumlah kalimatnya dari penulis novel [5].

1.2 Identifikasi Masalah

Dalam konteks penerapan *Natural Language Processing* (NLP) dengan algoritma *Support Vector Machines* (SVM) untuk klasifikasi genre novel bahasa Indonesia, berikut permasalahan yang dapat diidentifikasi:

1. Algoritma platform wattpad hanya menampilkan novel-novel dengan popularitas tinggi dari daftar genre.
2. Popularitas novel yang tidak memiliki genre hanya bergantung terhadap hashtag saja.
3. Pengambilan data genre yang terbatas dikarenakan menggunakan bahasa Indonesia.

1.3 Perumusan Masalah

Berdasarkan masalah yang ditemukan, maka dapat dirumuskan masalah-masalahnya adalah.

1. Bagaimana penggunaan *Natural Language Processing* dengan algoritma *Support Vector Machines* dapat digunakan untuk mengklasifikasikan genre

novel bahasa Indonesia, khususnya untuk menentukan persentase kalimat yang termasuk dalam genre romantis?

1.4 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan yang ingin dicapai, yaitu.

1. Mengembangkan model klasifikasi menggunakan algoritma Support Vector Machines (SVM) untuk menentukan persentase kalimat dalam novel yang termasuk dalam genre romantis.
2. Meningkatkan peluang penulis untuk mempromosikan karya tulis yang mereka buat dengan menggunakan potensi dari NLP dan SVM dalam menentukan genre novel.
3. Mengoptimalkan pengalaman membaca di era pergeseran minat pembaca dari novel fisik ke novel digital, khususnya melalui platform Wattpad, dengan menggunakan teknologi NLP dan SVM untuk membantu pengguna menavigasi dan menemukan karya sesuai dengan preferensi genre.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi signifikan terutama dalam:

1. Memberikan peluang untuk meningkatkan popularitas penulis baru.
2. Menjadikan penelitian ini sebagai perbandingan dengan peneliti lain dalam penggunaan NLP dan SVM di area karya tulis novel bahasa Indonesia.
3. Sebagai bahan evaluasi dalam penggunaan data novel dan data genre dalam klasifikasi.

1.6 Batasan Penelitian

Penelitian ini memiliki batasan-batasan tertentu yang perlu diperhatikan dengan tujuan untuk menghindari penyimpangan isu dan perbincangan terlalu luas serta memperhatikan keterbatasan informasi dan kapasitas pencipta, batasannya yang perlu diperhatikan adalah:

1. Data novel yang digunakan hanya dari platform wattpad saja.
2. Proses yang dilakukan hanya sampai model klasifikasi saja.

1.7 Sistematika Penulisan

Sistematika penulisan skripsi ini terbagi ke dalam beberapa bab yang secara garis besar terdiri dari:

BAB I PENDAHULUAN

Pada bab pendahuluan ini membahas mengenai latar belakang, identifikasi masalah, perumusan masalah, tujuan penelitian, manfaat penelitian, dan Batasan penelitian.

BAB II TINJAUAN PUSTAKA

Pada bab tinjauan Pustaka ini berisi kajian dari teori yang akan membantu dalam menyelesaikan permasalahan penelitian dan juga penelitian sebelumnya.

BAB III METODOLOGI PENELITIAN

Pada bab metodologi penelitian ini berisi penjelasan tahapan – tahapan penelitian, metode penelitian dan teknik – teknik yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini membahas dari hasil penelitian yang telah dilakukan secara detail.

BAB V PENUTUP

Pada bab ini menjelaskan kesimpulan untuk menjawab perumusan masalah secara terperinci dan rekomendasi perbaikan untuk kedepannya.