



**ANALISIS SENTIMEN ULASAN *GAME* HARRY POTTER: HOGWARTS
MYSTERY PADA SITUS GOOGLE PLAY MENGGUNAKAN *NAÏVE BAYES*
*CLASSIFIER***

SKRIPSI

PUTI AYU ANDHINI RAHMAN

1710512085

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI S-1 SISTEM INFORMASI

2021



**ANALISIS SENTIMEN ULASAN *GAME* HARRY POTTER: HOGWARTS
MYSTERY PADA SITUS GOOGLE PLAY MENGGUNAKAN *NAÏVE BAYES*
*CLASSIFIER***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh
Gelar Sarjana Komputer**

PUTI AYU ANDHINI RAHMAN

1710512085

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI S-1 SISTEM INFORMASI

2021

LEMBAR PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Puti Ayu Andhini Rahman

NIM 1710512085

Tanggal : 7 Juli 2021

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Depok, 1 Juli 2021

Yang menyatakan,



(Puti Ayu Andhini Rahman)

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK
KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Puti Ayu Andhini Rahman
NIM : 1710512085
Fakultas : Ilmu Komputer
Program Studi : S1 – Sistem Informasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non Eksklusif (*Non-exclusive Royalti Free Right*) atas karya ilmiah saya yang berjudul:

ANALISIS SENTIMEN ULASAN *GAME* HARRY POTTER: HOGWARTS MYSTERY PADA SITUS GOOGLE PLAY MENGGUNAKAN *NAÏVE BAYES CLASSIFIER*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap menyantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok

Pada tanggal : 1 Juli 2021

Yang menyatakan,



(Puti Ayu Andhini Rahman)

LEMBAR PENGESAHAN

Skripsi diajukan oleh:

Nama : Puti Ayu Andhini Rahman
NIM : 1710512085
Program Studi : S-1 Sistem Informasi
Judul Skripsi : Analisis Sentimen Ulasan *Game* Harry Potter: Hogwarts Mystery pada Situs Google Play Menggunakan *Naïve Bayes Classifier*

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S-1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Anita Muliawati, S.Kom., MTL.

Penguji I



Ika Nurlaili Isnainivah, S.Kom., MMSI.

Penguji II



Dr. Ermatita, M.Kom.

Pembimbing I



Helena Nurramdhani Irmada, S.Pd., M.Kom.

Pembimbing II



Dr. Ermatita, M.Kom.

Dekan



Ati Zaidiah, S.Kom., MTL.

Ketua Program Studi



Ditetapkan di : Jakarta
Tanggal Ujian : 14 Juli 2021

LEMBAR PERSETUJUAN

Dengan ini menyatakan bahwa proposal berikut:

Nama : Puti Ayu Andhini Rahman

NIM : 1710512085

Program Studi : S1 – Sistem Informasi

Judul : Analisis Sentimen Ulasan *Game* Harry Potter: Hogwarts Mystery pada Situs Google Play Menggunakan *Naïve Bayes Classifier*

Sebagai bagian persyaratan yang diperlukan untuk mengikuti ujian Sidang Proposal Tugas Akhir/Skripsi pada Program Studi S1- Sistem Informasi Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Menyetujui,

Dosen Pembimbing 1



(Dr. Ermatita, M.Kom.)

Dosen Pembimbing 2



(Helena Nurramdhani I., S.Pd., M.Kom.)

Mengetahui,

Ketua Program Studi



(Ati Zaidiah., S.Kom., MTI)

Ditetapkan : Jakarta

Tanggal Persetujuan : 1 Juli 2021

**ANALISIS SENTIMEN ULASAN *GAME* HARRY POTTER: HOGWARTS
MYSTERY PADA SITUS GOOGLE PLAY MENGGUNAKAN *NAÏVE BAYES*
*CLASSIFIER***

Puti Ayu Andhini Rahman

ABSTRAK

Salah satu dampak dari pandemi Covid-19 adalah semua orang harus berada di rumah untuk meminimalisir penyebaran virus dan mereka merasa jenuh karena taman rekreasi atau bermain tutup, untuk menghilangkan rasa jenuh tersebut banyak dari mereka yang mencari hiburan melalui permainan video baik *online* maupun *offline*. Harry Potter: Hogwarts Mystery adalah salah satu permainan video *online* bermain peran yang berlatarkan cerita Harry Potter berdasarkan seri novel JK Rowling. Permainan ini masuk kedalam pilihan editor di Google Play Store, yang dikembangkan oleh Jam City dengan lisensi dari Portkey Games, yang dirilis pada 25 April 2018 untuk perangkat Android dan Ios. Dalam pengembangan permainan ini, tentunya ulasan dari pengguna merupakan satu dari banyak hal penting untuk dijadikan pertimbangan oleh developer. Ulasan pengguna yang tersimpan dalam basis data Google Play Store dapat dijadikan bahan untuk penelitian analisis pengguna. Ulasan dari pengguna permainan ini umumnya berisi kesan positif atau kritik yang bersifat negatif, dimana hal ini tentu berpengaruh terhadap rating sebuah permainan agar permainan dapat terus berkembang. Untuk memantau ulasan-ulasan tersebut, dibutuhkan suatu metode yang dapat memantau secara garis besar sentimen pengguna, metode yang dilakukan pada penelitian ini menyangkut kegiatan pengumpulan data menggunakan *web scrapper*, melalui *preprocessing* data, pembobotan kata, dan teknik yang digunakan untuk klasifikasi sentimen menggunakan *Naïve Bayes Classifier* yang dapat mengkategorikan ulasan pengguna tanpa harus melakukannya secara manual. Tujuan dari penelitian ini adalah untuk melakukan klasifikasi terhadap ulasan tersebut dan mendapatkan informasi yang berguna bagi pihak yang berkepentingan. Hasil dari klasifikasi diharapkan mampu mengkategorikan kelompok sentimen pengguna game Harry Potter: Hogwarts Mystery dengan tingkat akurasi yang cukup tinggi.

Kata kunci: Analisis Sentimen, *Google Play*, *Game*, *Naïve Bayes Classifier*

**SENTIMENT ANALYSIS OF HARRY POTTER: HOGWARTS MYSTERY
GAME REVIEW ON GOOGLE PLAY SITE USING *NAÏVE BAYES*
CLASSIFIER**

Puti Ayu Andhini Rahman

ABSTRACT

One of the impacts of Covid-19 pandemic is everyone must be at home to minimize the spread of the virus and they feel bored because recreational park or play places are closed, to eliminate this feeling of saturation, many of them are looking for entertainment through video games both online and offline. Harry Potter: Hogwarts Mystery is an online role-playing video game set in the Harry Potter story based on the JK Rowling novel series. This game is included in the editor's choice on the Google Play Store, developed by Jam City with a license from Portkey Games, which was released on April 25th, 2018 for Android and iOS devices. In developing this game, of course, user reviews are one of the many important things for developers to consider. User reviews stored in the Google Play Store database can be used as research material for user analysis. Reviews from game users containing positive impressions or negative criticism, which certainly affects the rating of game, in order that the game can continue to develop. To monitor these reviews, a method is needed that can outline user sentiment, the method used in this study involves data collection activities using a web scraper, through data preprocessing, word weighting, and techniques used for sentiment classification using Naïve Bayes. A classifier that can categorize user reviews without having to do it manually. The purpose of this study is to classify these reviews and obtain useful information for interested parties. The results of the classification are expected to be able to categorize the sentiment groups of Harry Potter: Hogwarts Mystery users with a fairly high level of accuracy.

Keywords: Sentiment Analysis, Google Play, Game, Naïve Bayes Classifier

KATA PENGANTAR

Puji dan syukur kepada Allah SWT berkat rahmat dan karunia-Nya, penulis dapat mengerjakan dan menyelesaikan Proposal Tugas Akhir. Penulis menyadari bahwa dalam penyusunan Proposal Tugas Akhir ini, penulis menerima banyak bantuan dan dukungan dari berbagai pihak. Untuk itu, pada kesempatan kali ini penulis ingin mengucapkan terima kasih kepada:

1. Ibu Dr. Ermatita, M.Kom. selaku Dekan Fakultas Ilmu Komputer dan sebagai Dosen Pembimbing I.
2. Ibu Anita Muliawati, S.Kom., MTI. selaku Kepala Jurusan Program Studi Sistem Informasi.
3. Ibu Ati Zaidiah, S.Kom, MTI. selaku Ketua Program Studi S1 Sistem Informasi.
4. Ibu Helena Nurramdhani Irmanda., S.Pd., M.Kom selaku Dosen Pembimbing II.
5. Orangtua dan keluarga besar yang telah memberikan dukungan baik secara moril dan materil.
6. Teman baik penulis yang telah membantu penulis dalam memberikan saran dan dukungan dalam penulisan laporan tugas akhir ini.

Akhir kata, penulis menyadari bahwa Tugas Akhir ini masih banyak kekurangan dan pengembangan, untuk itu penulis mohon kritik dan saran yang membangun agar dapat lebih baik lagi, semoga Tugas Akhir ini dapat bermanfaat bagi yang membutuhkan, terima kasih.

Depok, 1 Juli 2021



Penulis

Puti Ayu Andhini Rahman

DAFTAR ISI

LEMBAR PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	iv
LEMBAR PENGESAHAN	v
LEMBAR PERSETUJUAN	v
ABSTRAK	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
DAFTAR SIMBOL	xv
DAFTAR LAMPIRAN	xvi
BAB I.....	7
PENDAHULUAN	7
1.1 Latar Belakang Masalah	7
2.1 Rumusan Masalah.....	11
2.2 Batasan Masalah	11
2.3 Tujuan Penelitian.....	11
2.4 Manfaat Penelitian.....	11
2.5 Sistematika Penulisan	12
BAB II TINJAUAN PUSTAKA.....	13
2.1 Analisis Sentimen	13
2.2 Game	13
2.2.1 RPG (Role Playing Game).....	14
2.3 Harry Potter: Hogwarts Mystery	14
2.4 Data Mining.....	15
2.4.1 Text Mining	15
2.5 Preprocessing	16
2.5.1 Data Cleaning.....	16
2.5.2 Tokenization	17
2.5.3 Filtering.....	17
2.5.4 Case Folding	17
2.5.5 Stemming.....	17

2.6	Pembobotan Kata.....	18
2.6.1	TF-IDF	18
2.7	Klasifikasi	19
2.8	Naïve Bayes Classifier.....	19
2.9	Penelitian Sebelumnya.....	21
BAB III METODOLOGI PENELITIAN		25
3.1	Tahapan Penelitian	25
3.1.1	Menentukan Topik.....	26
3.1.2	Identifikasi Masalah	26
3.1.3	Studi Literatur.....	26
3.1.4	Pengumpulan Data Ulasan.....	26
3.1.5	Praproses	27
3.1.6	Pembobotan TF-IDF	27
3.1.7	Klasifikasi NBC	28
3.1.8	Evaluasi.....	29
3.1.9	Interpretasi Hasil	30
3.2	Kegiatan Penelitian.....	30
3.2.1	Populasi dan Sampel Penelitian.....	30
3.2.2	Jenis dan Sumber Data.....	31
3.2.3	Variabel Penelitian.....	31
3.3	Metode Pengumpulan Data.....	32
3.4	Metode Analisis Data	32
3.5	Alat dan Bahan Penelitian.....	32
BAB IV		34
HASIL DAN PEMBAHASAN		34
4.1	Pengumpulan Data Ulasan.....	34
4.2	Praproses	37
4.2.1	Terjemah ulasan Bahasa asing	37
4.2.2	Pelabelan data	38
4.2.3	Data Cleaning	39
4.2.3.1	Menghapus variable yang tidak perlu.....	40
4.2.4	Tokenization	41
4.2.5	Filtering(stopwords).....	41
4.2.6	Case Folding	42
4.2.7	Stemming.....	43
4.3	Pembobotan TF-IDF.....	44

4.4 Klasifikasi NBC	45
4.4.1 Pembagian Data Latih dan Data Uji	46
4.4.2 Pembentukan Model Naïve bayes	47
4.5 Evaluasi Algoritma Naïve Bayes	50
4.5.1 Confusion Matrix	50
4.5.2 Recall	52
4.5.3 Precision	53
4.5.4 F1-Score	55
4.5.5 Accuracy	56
4.6 Analisis Interpretasi Hasil dan Visualisasi	57
4.6.1 Analisis Interpretasi Hasil	58
4.6.2 Visualisasi <i>Word Cloud</i>	58
4.6.2.1 Visualisasi <i>Word Cloud</i> pada Kelas Positif	59
4.6.2.2 Visualisasi <i>Word Cloud</i> pada Kelas Negatif	60
BAB V	62
5.1 Kesimpulan	62
5.2 Saran	62
DAFTAR PUSTAKA	64
RIWAYAT HIDUP	66
LAMPIRAN	67

DAFTAR GAMBAR

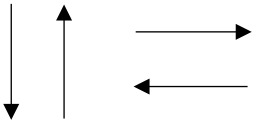
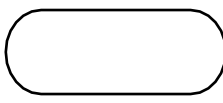

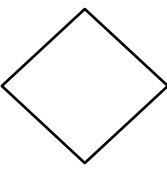
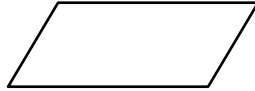
Gambar 1 Global Games Market 2021	8
Gambar 2 Penjualan konsol bersumber dari artikel terbitan Aranca.....	9
Gambar 3 Tahap Penelitian	25
Gambar 4 Rumus Evaluasi.....	30
Gambar 5 Tampilan finder tool dataminer.io	34
Gambar 6 Tampilan scraping data luaran pada website.....	35
Gambar 7 Scraping data ulasan menggunakan dataminer.io	35
Gambar 8 Variabel data ulasan	36
Gambar 9 Hasil scraping data pada dataminer.io.....	37
Gambar 10 Tampilan rating	39
Gambar 11 Proses pelabelan data.....	39
Gambar 12 Tampilan proses data cleaning	40
Gambar 13 Tampilan proses penghapusan variabel tidak penting.....	40
Gambar 14 Kalimat sebelum tokenisasi.....	41
Gambar 15 Proses tokenisasi	41
Gambar 16 Tfidf Vectorizer pada Python	44
Gambar 17 Visualisasi Word Cloud Kelas Positif dan Negatif	58
Gambar 18 Word Cloud Kelas Positif.....	59
Gambar 19 Word Cloud Kelas Negatif	60

DAFTAR TABEL

Tabel 1 Penelitian terdahulu.....	20
Tabel 2 Contoh pelabelan sentimen	27
Tabel 3 Simulasi perhitungan label pada sentimen.....	27
Tabel 4 Rumus Confusion Matrix.....	28
Tabel 5 Proses terjemah ulasan.....	38
Tabel 6 Proses Filtering (Stopwords).....	42
Tabel 7 Proses Case Folding.....	43
Tabel 8 Proses Stemming.....	43
Tabel 9 Urutan bobot kata keseluruhan	45
Tabel 10 Pembagian data latih dan data uji	46
Tabel 11 Normalisasi Data Sampel.....	48
Tabel 12 Contoh Dataset untuk simulasi	49
Tabel 13 Perhitungan Probabilitas Simulasi	47
Tabel 14 Confusion Matrix rangkaian pertama	50
Tabel 15 Confusion Matrix rangkaian kedua.....	51
Tabel 16 Recall pada rangkaian pertama	52
Tabel 17 Recall pada rangkaian kedua.....	52
Tabel 18 Precision pada rangkaian pertama.....	53
Tabel 19 Precision pada rangkaian kedua.....	54
Tabel 20 F1-Score pada rangkaian pertama.....	55
Tabel 21 F1-Score pada rangkaian kedua	55
Tabel 22 Accuracy pada rangkaian pertama	56
Tabel 23 Accuracy pada rangkaian kedua	57
Tabel 24 Jadwal Penelitian.....	77

DAFTAR SIMBOL

a.) *Flowchart*

Nama	Simbol	Deskripsi
<i>Flow</i>		<i>Flow</i> merupakan penunjuk arah aliran antar proses.
<i>Terminator</i>		<i>Terminator</i> merupakan simbol untuk memulai atau mengakhiri sebuah program.
<i>Process</i>		<i>Process</i> merupakan simbol yang menunjukkan kegiatan yang dilakukan komputer.
<i>Decision</i>		<i>Decision</i> merupakan simbol proses/langkah dimana perlu adanya proses keputusan akan kondisi tertentu.
<i>Input – Output</i>		Simbol <i>Input – Output</i> menunjukkan proses keluaran dan masukan data tergantung dengan jenis peralatannya.

DAFTAR LAMPIRAN

Lampiran 1 Import dan Pendefinisian Data	67
Lampiran 2 Persiapan Library untuk Preprocessing	67
Lampiran 3 Pengubahan Kolom Ulasan Menjadi List	68
Lampiran 4 Preprocessing	68
Lampiran 5 Pembagian Data Training dan Data Testing	73
Lampiran 6 Proses Pemodelan Klasifikasi	74
Lampiran 7 Visualisasi Menggunakan Wordcloud	75
Lampiran 8 Jadwal Penelitian	77