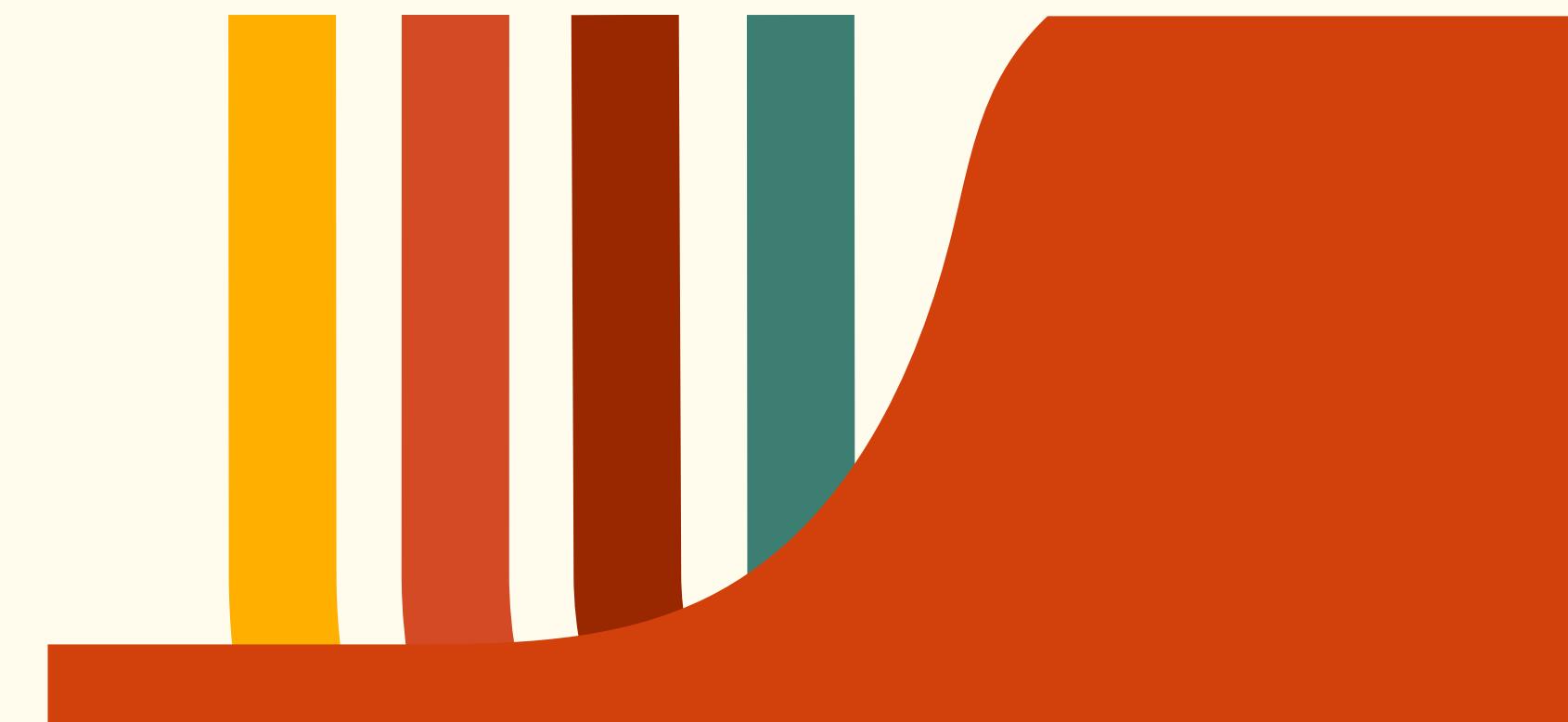


Penggunaan Topic Modelling untuk Analisis Data Ulasan Pemain Game Honkai Star Rail

*Mata Kuliah Data Mining
Jurusan Informatika*

Disusun Oleh:
Annisa Mardhatilaah (3337220029)
Rafi Kyandra Atharizqi (3337220052)

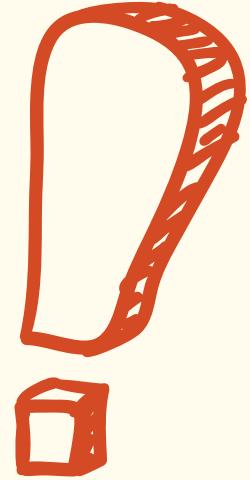




HONKAI STAR RAIL
TOPIC MODELLING
PROCESS
DATA
EDA
DATA PREPROCESSING
WORD CLOUD
COHERENCE SCORE ANALYSIS
INSIGHT FROM TOPIC MODELLING



Honkai Star Rail



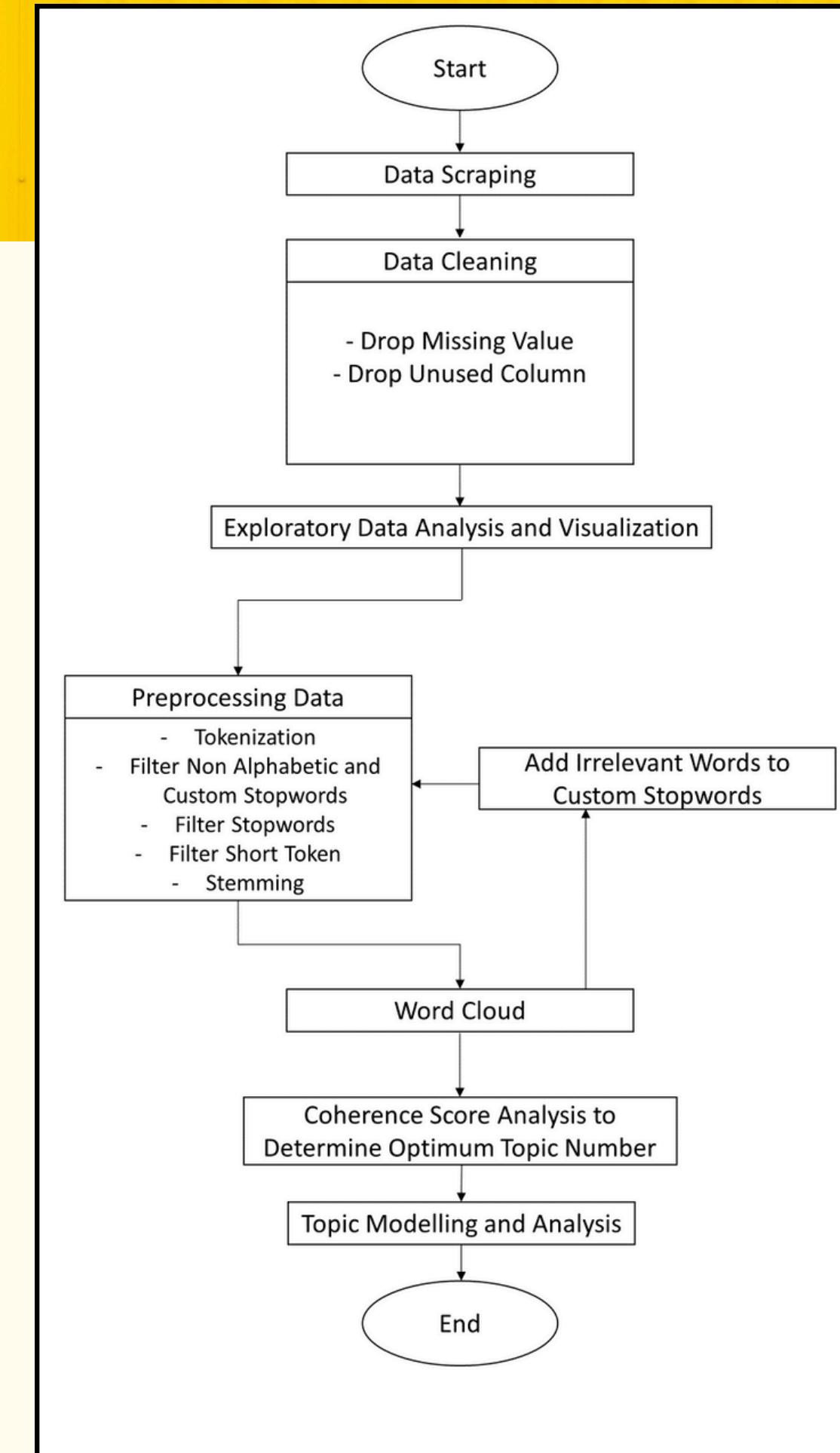
Honkai: Star Rail adalah sebuah permainan video RPG berbasis giliran yang dikembangkan oleh miHoYo, yang dikenal melalui judul-judul sukses seperti Genshin Impact dan Honkai Impact 3rd. Dirilis pada 26 April 2023 untuk PC, iOS, dan Android, permainan ini menggabungkan eksplorasi dunia yang luas dengan pertarungan strategis berbasis giliran.

Topic

Modelling

- Teknik dalam pemrosesan bahasa alami (NLP) yang digunakan untuk mengidentifikasi topik-topik tersembunyi dalam kumpulan dokumen teks.
- Algoritma seperti Latent Dirichlet Allocation (LDA) mengasumsikan bahwa setiap dokumen adalah campuran dari beberapa topik, dan setiap topik adalah campuran dari sejumlah kata.
- Dengan menganalisis pola kemunculan kata, topic modelling dapat mengungkap struktur tematik dalam dokumen, memfasilitasi analisis konten, rekomendasi konten, dan penelitian akademik.
- Teknik ini memungkinkan pengelompokan informasi teks yang besar dan kompleks menjadi topik-topik yang lebih mudah dipahami dan dianalisis.

Process



Nama Proses	Penjelasan
Data Gathering	Menggunakan Web Scraping untuk ulasan game di Google Play Store. Hasil berupa file CSV
Data Cleaning	Menghilangkan data yang memiliki nilai null dan column data yang tidak dipakai (userName dan at)
EDA	Memvisualisasikan data yang sudah ada untuk menggali wawasan
Data Preprocessing	<p>Menyiapkan data isi ulasan (content) untuk proses selanjutnya.</p> <ul style="list-style-type: none"> • Tokenization: Data review akan dipisah dipecah menjadi kata-kata • Filter Non Alphabetic: Menghilangkan token yang tidak berada dalam alphabet • Filter Custom Stopwords: Menghapus token yang termasuk Stopwords tambahan. Stopwords adalah kata-kata yang tidak diinginkan muncul pada proses selanjutnya, biasanya berisi kata yang tidak relevan. • Filter Stopwords: Menghapus token yang termasuk Stopwords pada library Sastrawi • Filter Short Token: Menghapus token yang pendek (satu karakter) • Stemming: Mengubah setiap token menjadi bentuk dasarnya, contoh Mengajar akan menjadi ajar.

Word Cloud	Memvisualisasikan kata-kata yang paling sering muncul di setiap kategori menggunakan teknik Word Cloud.
Coherence Score Analysis to Determine Optimum Topic Numbers	Mencari jumlah topic dengan rata - rata coherence score yang paling tinggi untuk digunakan pada Topic Modelling
Topic Modelling and Analysis	Proses penggunaan algoritma dan analisanya

DATA

Google Play Store

SUMBER

398

ROWS

Web Scraping

TEKNIK

4*

COLUMN



DATA

userName

Berisi dengan nama pengguna yang membuat review/ulasan

score

Representasi skor bintang yang diberikan pengguna.
Nominalnya dari 1 sampai 5

at

Waktu ketika ulasan tersebut dibuat

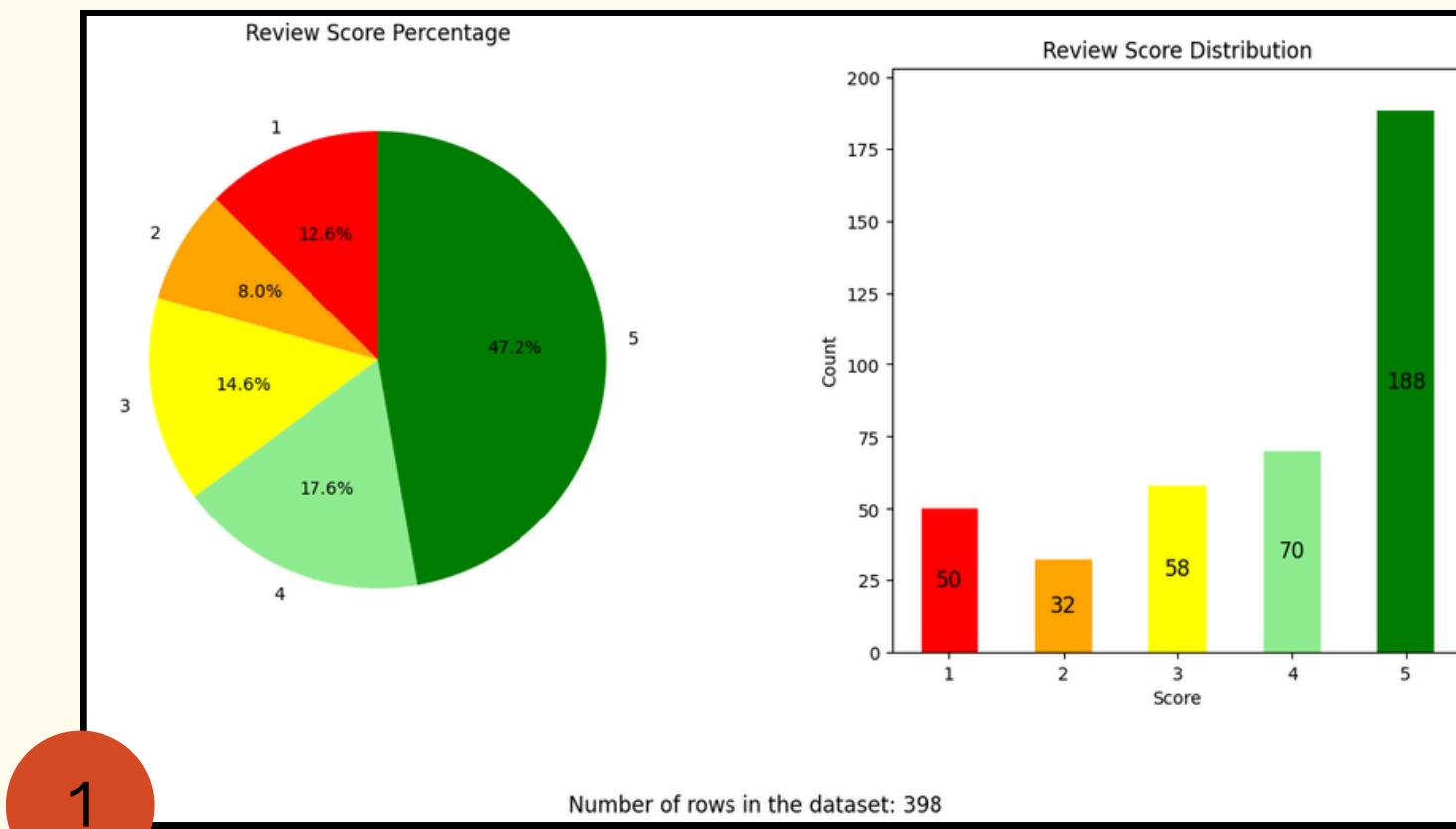
content

Isi ulasan

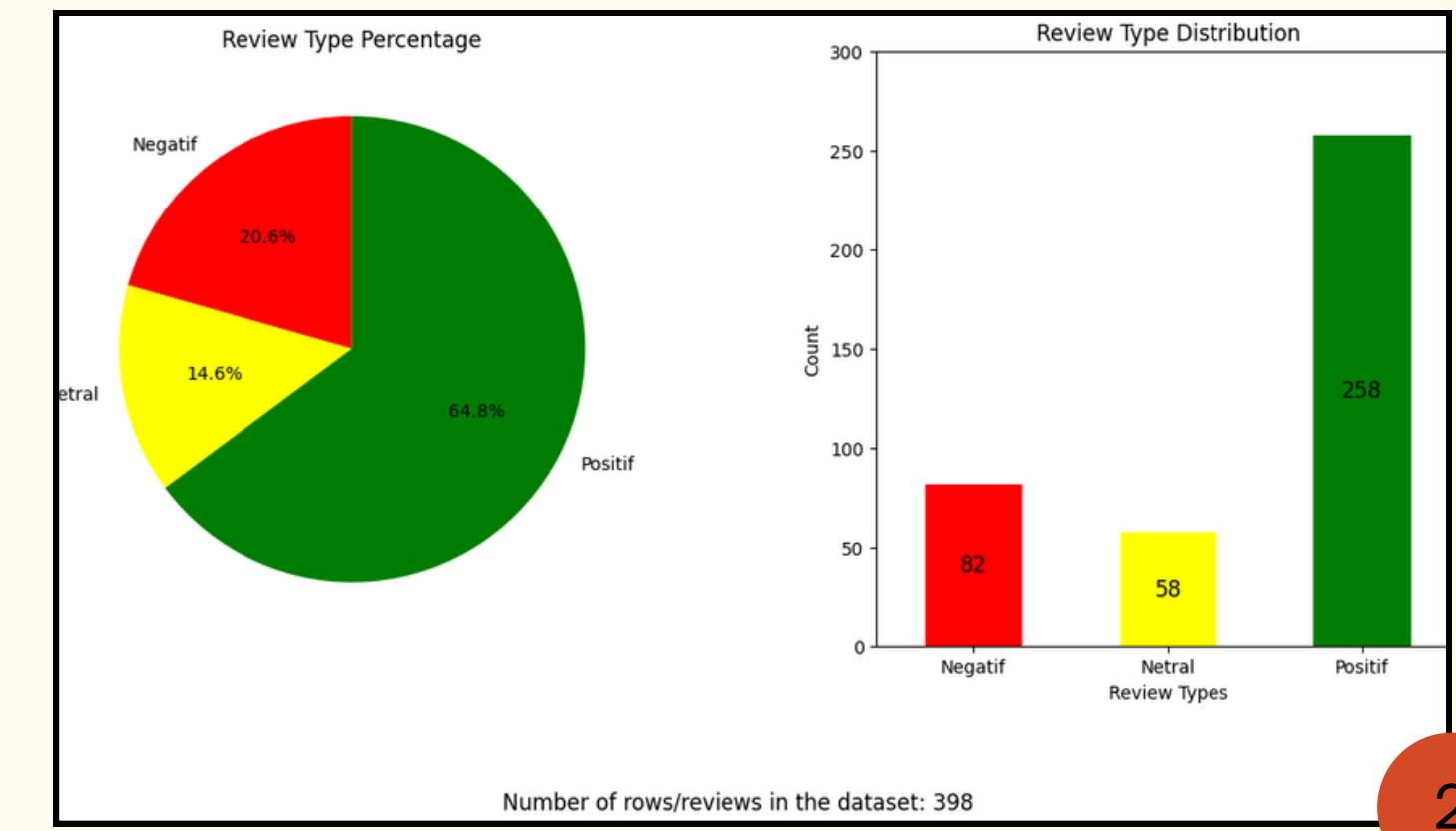
type

Kolom yang ditambahkan pada CSV. Pembagian score menjadi 3 tipe, yaitu

- Negatif (1-2)
- Netral (3)
- Positif (4-5)

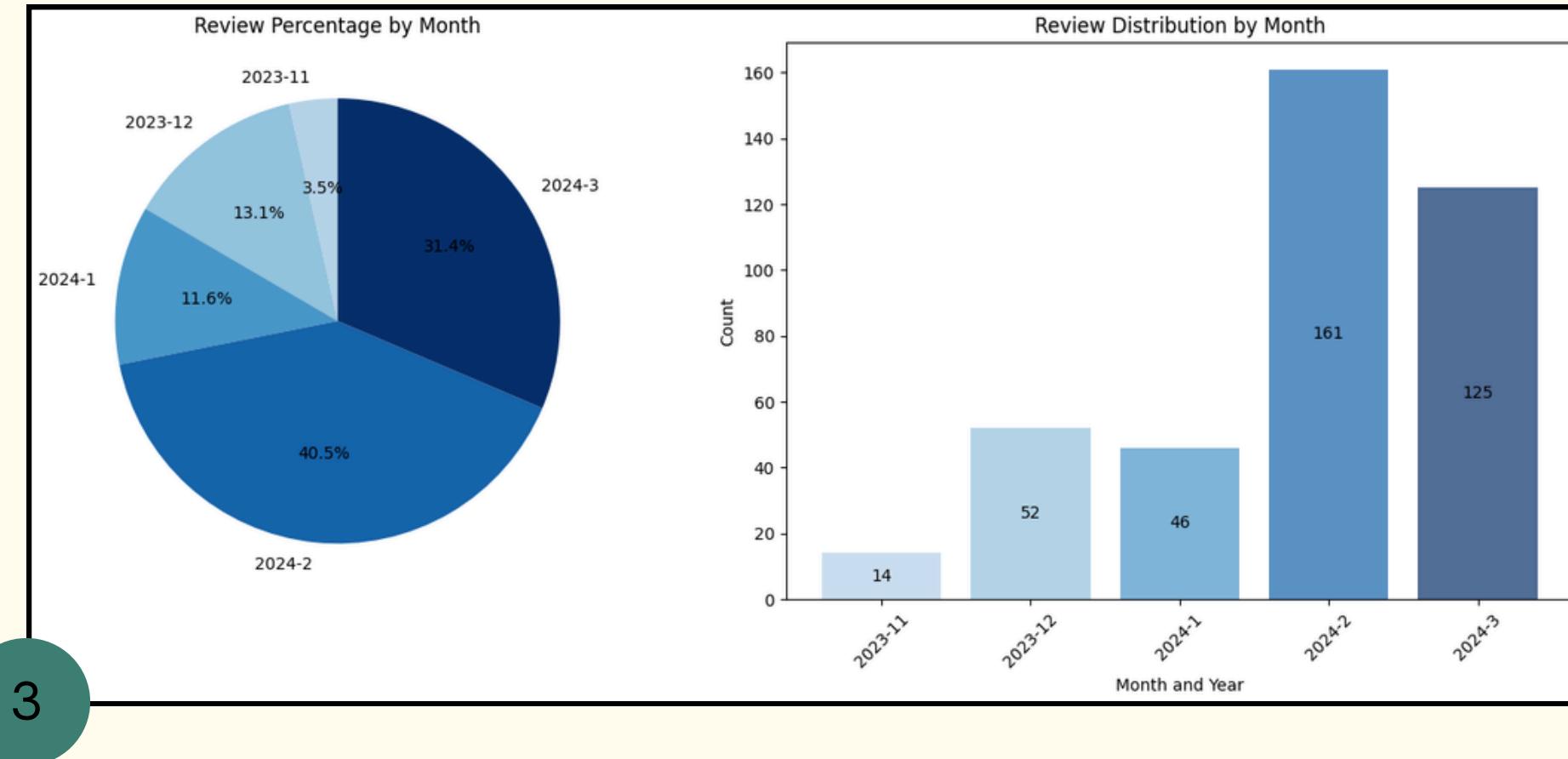


1

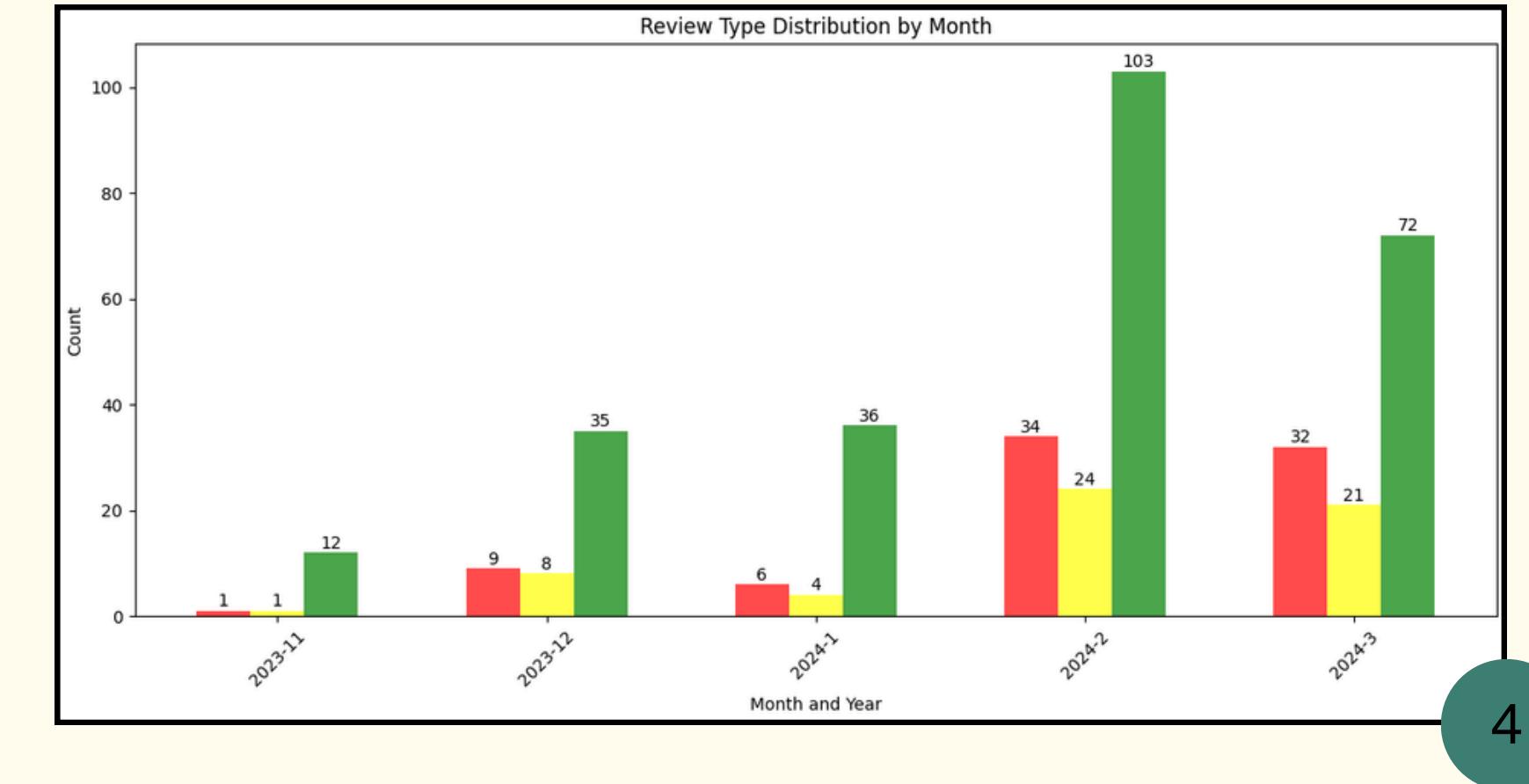


2

EDA



3



4

1

Terlihat bahwa data dengan skor 5 memiliki jumlah paling banyak. Selisihnya pun cukup besar kepada skor terbanyak kedua (Skor 4), yaitu 29.6%, hampir mencapai 30%. Hal ini menunjukkan bahwa mayoritas pengguna sangat menyukai atau puas dengan game hingga menganggap game pantas dinilai dengan bintang atau skor 5, skor yang paling besar.

2

Hasil nya mirip dengan poin sebelumnya, dimana tipe Positif sangat dominan. Akan tetapi, dapat kita amati bahwa tipe Negatif melebihi tipe Netral, menandai bahwa masih banyak pengguna yang merasa sangat tidak puas, hal ini dapat menjadi sebuah pertimbangan kebijakan pengembangan.

3

Periode ulasan yang ada pada CSV adalah dari November 2023 sampai dengan Maret 2024. Pada Bulan Februari 2024, user paling banyak memberikan ulasan. Hal itu dikarenakan pada bulan tersebut game Honkai Star Rail mendapatkan update Version 2.0 yang menambah wilayah baru pada game.

4

Hasil pengamatan menunjukkan bahwa distribusi tipe review apabila kita pecah menjadi setiap bulan mirip dengan tabel tipe review keseluruhan, dimana tipe Positif mendominasi dan tipe Negatif memiliki kecenderungan untuk lebih banyak dari netral.

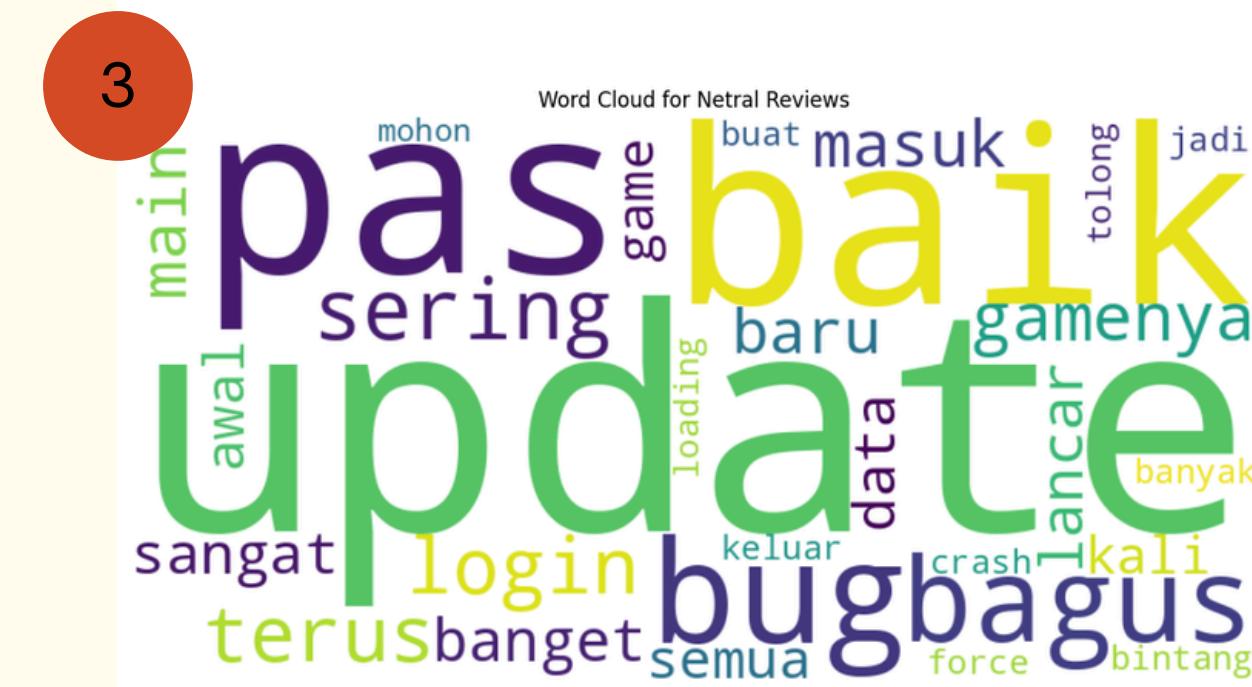
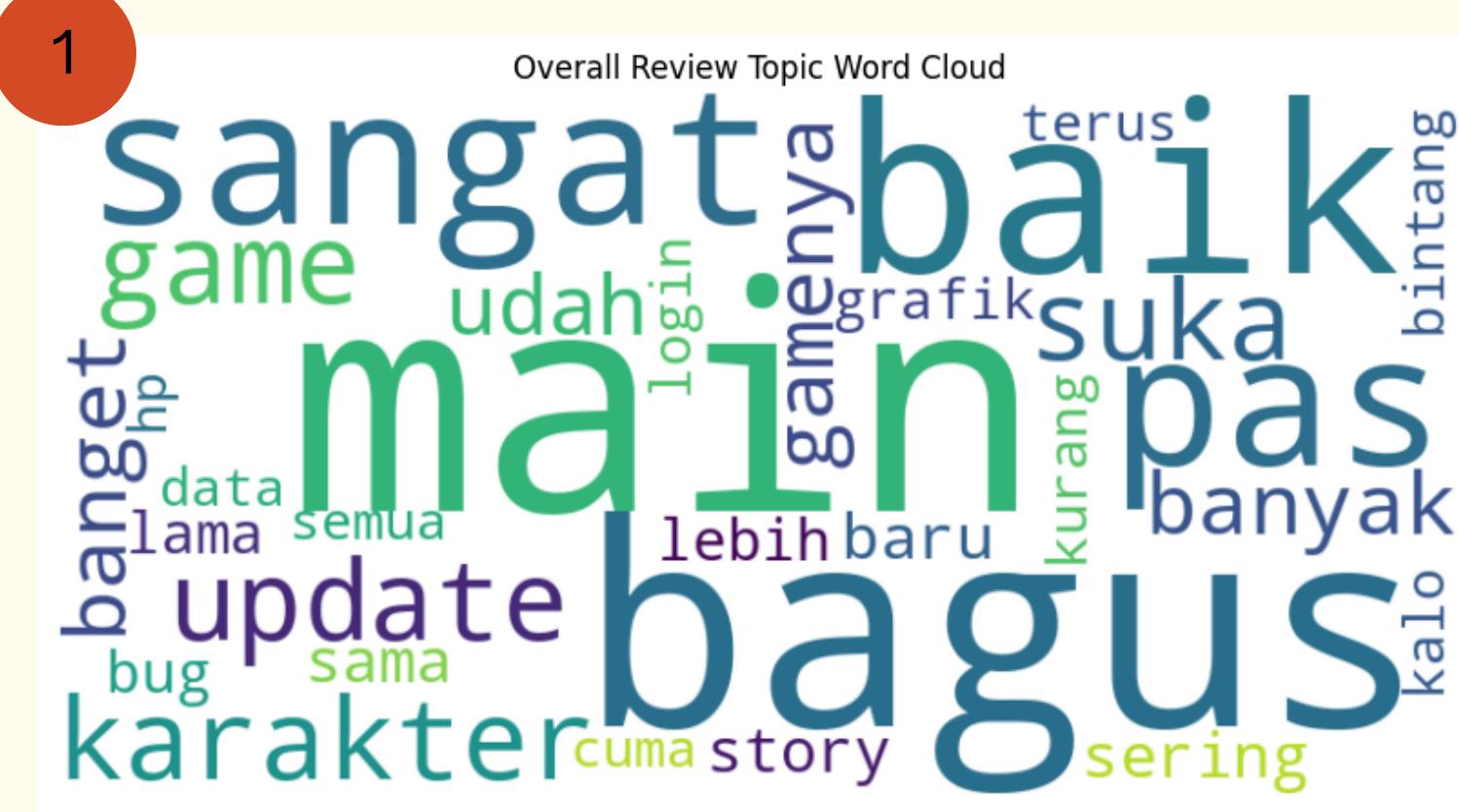
Data Preprocessing

Menyiapkan data isi ulasan (content) untuk proses selanjutnya.

- Tokenization: Data review akan dipisah dipecah menjadi kata-kata
Filter Non Alphabetic:
Menghilangkan token yang tidak berada dalam alphabet
- Filter Custom Stopwords: Menghapus token yang termasuk Stopwords tambahan.
Stopwords adalah kata-kata yang tidak diinginkan muncul pada proses selanjutnya,
biasanya berisi kata yang tidak relevan.
- Filter Stopwords: Menghapus token yang termasuk Stopwords pada library Sastrawi
- Filter Short Token: Menghapus token yang pendek (satu karakter)
- Stemming: Mengubah setiap token menjadi bentuk dasarnya, contoh Mengajar akan
menjadi ajar.



Word Cloud



1

Dari Word Cloud untuk data keseluruhan, diketahui bahwa kata-kata positif seperti bagus, baik, dan pas paling sering muncul. Hal ini sesuai dengan jumlah data positif yang lebih banyak dari tipe data lainnya. Kata-kata lain seperti karakter, story, grafik juga banyak disebutkan oleh pengguna, hal itu menjadi fitur-fitur yang harus dikedepankan oleh game.

2

Disimpulkan bahwa terdapat banyaknya update secara konstan dan memiliki ukuran yang besar pada setiap update. Selain itu besarnya RAM yang digunakan saat menjalankan apliksi ini menjadi salah satu masalah yang dihadapi para pengguna yang hanya menggunakan device mobile seadanya. Hal ini menyebabkan para pemain tidak dapat masuk kedalam game ini. Akan lebih baik jika developer melakukan optimasi lebih lanjut untuk mengatasi masalah yang dihadapi pengguna mobile.

3

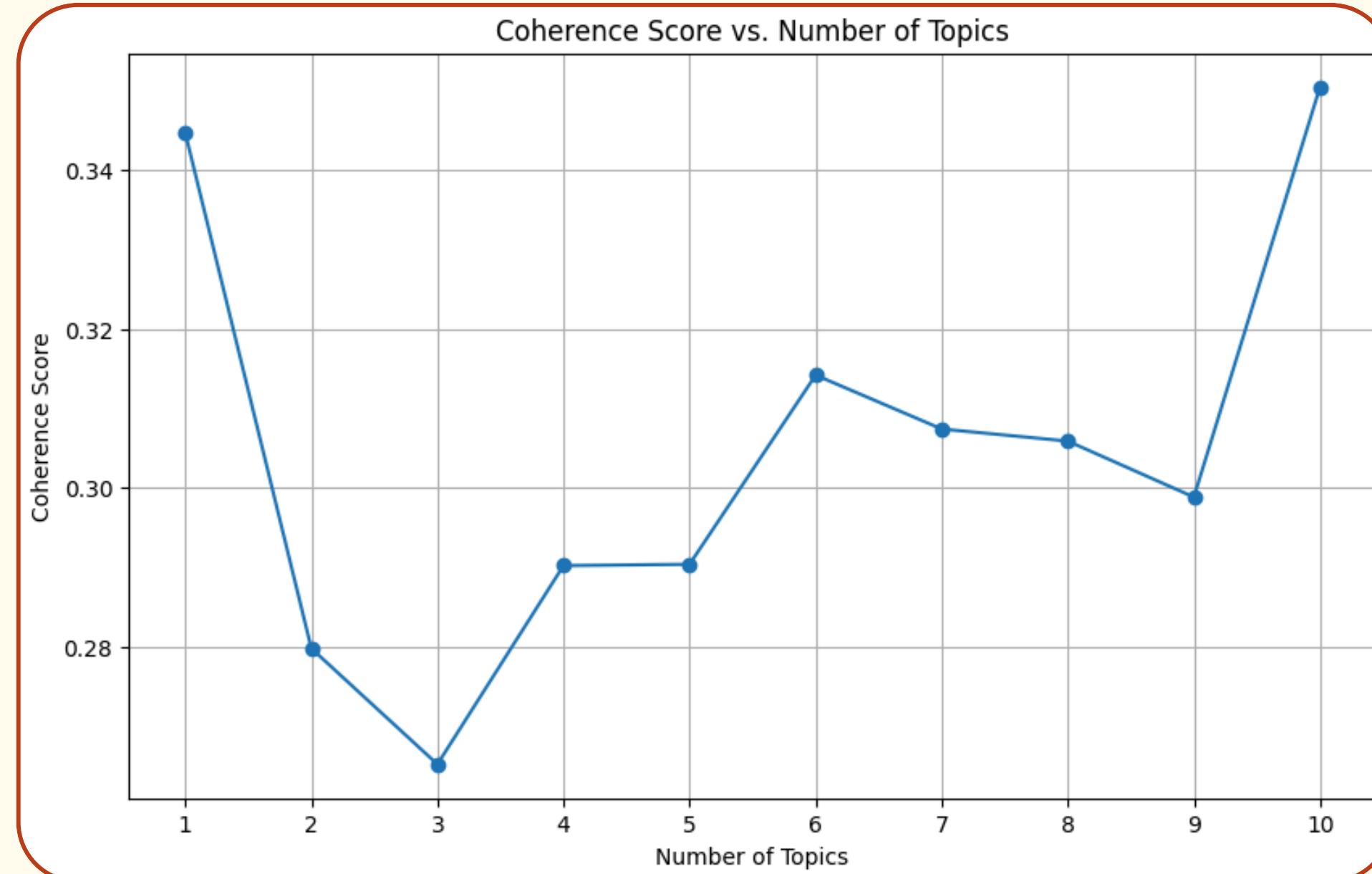
Pada kumpulan dari World Cloud dengan tipe review netral tampaknya para pemain cukup puas dari apa yang dihadirkan oleh Honkai: Star Rail ini. Walaupun sepertinya masih ada keluhan terkait besarnya ukuran pada setiap updatenya. Adanya bug dalam game juga mengganggu pengalaman para pemain walaupun, bug muncul sangat jarang. Tapi, sayangnya bug minor tersebut tidak pernah diperbaiki oleh developer.

4

Kumpulan World Cloud dari tipe review positif memberikan bukti kalau banyak pemain yang merasa sangat puas terhadap konten, karakter, visual, gameplay, ataupun cerita yang dihadirkan dalam game Honkai: Star rail ini. Para pemain dari tipe review ini tampaknya tidak terlalu memikirkan bug minor yang ada dalam game. Berbeda dari kedua tipe review sebelumnya disini para pemain nampaknya tidak ada masalah saat mencoba masuk ke dalam game.

Coherence Score Analysis

Coherence Score dengan nilai tertinggi didapatkan dengan jumlah topic 10, oleh karena itu saat Topic Modelling akan digunakan 10 topic.



Insight from Topic Modeling



Berdasarkan
Skor



Berdasarkan
Jenis



Berdasarkan Skor

*1
game sering forced closed, sulitnya mengumpulkan item yang diinginkan, update yang terlalu besar

*2
sinyal tidak stabil, sering ngefreeze saat di cutscene, dan game sering forced closed

*3
adanya bug pada game yang mempengaruhi pengalaman pemain, terjadinya crash pada pemain, pengalaman bermain yang kurang lancar

*4
dialog pada cerita terlalu panjang, puas dengan animasi ultimate karakter, adanya ruang untuk penambahan dari karakter yang dapat dimainkan

*5
pemain merasa puas terhadap cerita yang gihadirkan, desain karakter yang bagus menjadi alasan utama pemain, banyak pemain yang merasa puas terhadap hadiah dalam game yang diberikan



Berdasarkan jenis

negatif

netral

positif

Para pemain mengeluh kalau optimasi game Honkai: Star Rail harus ditingkatkan lagi. Selain itu, pemain mengeluh kalau adanya beberapa musuh yang terlalu sukit untuk dikalahkan.

Mirip dengan yang mereview negatif, terdapat beberapa keluhan yang mirip seperti review negatif. meskipun begitu mereka nampaknya cukup puas dari apa yang telah disajikan oleh game ini.

Berbeda dari kedua tipe review lainnya. para pemain yang memberikan review positif sepertinya tidak ada masalah yang dikeluhkan dan mereka merasa sangat puas dari yang game ini berikan.



Sekian dan
Terima kasih

