

## 1. Pendahuluan

### Latar Belakang

Graphics Processing Unit (GPU) merupakan salah satu komponen pada komputer yang bertugas menangani kebutuhan grafis pada komputer tersebut. Secara umum, GPU hampir mirip dengan Central Processing Unit (CPU), namun GPU memiliki lebih banyak inti dan tugas khusus dalam membantu CPU dalam menangani data dalam bentuk grafik seperti merender grafik, video, dan animasi. Namun, seiring dengan perkembangan GPU, penggunaannya menjadi lebih luas dan dapat menangani lebih banyak aplikasi, termasuk beberapa yang sedang dikembangkan. Perkembangan dari GPU tersebut menjadikan komponen ini menjadi salah satu komponen utama bagi para editor video, animator, gamer, dan pekerjaan lainnya yang terkait dengan grafik komputer. Nvidia merupakan salah satu vendor yang memproduksi beragam GPU sesuai dengan kebutuhan pengguna. Seri yang terlaris saat ini adalah seri Nvidia RTX [1], yang mana dibantu oleh kondisi Work from Home (WFH) dan pekerjaan remote lainnya yang membutuhkan komponen grafik dengan performa tinggi. Nvidia juga menggunakan sosial media dalam mengembangkan bisnis mereka. Twitter adalah salah satu pilihan Nvidia yang digunakan sebagai media pemasaran dan media interaksi dengan pengguna.

Dengan mempertimbangkan *tweet*, pengguna dapat membuat referensi seri GPU Nvidia yang sesuai dengan kebutuhan mereka dengan menyesuaikan berbagai aspek kategori seperti harga, performa, dan ketersediaan dari produk tersebut. Perusahaan dapat melakukan analisis pada *tweet* tersebut sebagai acuan kepuasan pengguna terhadap produk mereka, proses ini dinamakan *multi-aspect sentiment analysis*.

Banyak penelitian mengenai sentimen analisis yang telah dilakukan sebelumnya, terutama menggunakan Naive Bayes Classifier, contohnya penelitian yang dilakukan oleh Bayhaqy [2] pada tahun 2018, membahas bagaimana mengidentifikasi *tweet* dengan polaritas positif dan negatif. Para peneliti menggunakan dataset dari e-commerce Tokopedia dan Bukalapak dan membandingkan 3 classifier yaitu Naive Bayes, Decision Tree, dan K-Nearest Neighbour. Hasil akurasi tertinggi diperoleh dengan menggunakan classifier Decision Tree yaitu 80%, diikuti K-NN 78%, dan Naive Bayes 77%. Sentimen analisis juga banyak digunakan untuk menganalisis suatu produk seperti yang dilakukan oleh Haque [3] yang berfokus pada produk telepon genggam, elektronik, dan instrumen musik yang tersedia di e-commerce Amazon.

Berdasarkan penelitian sebelumnya mengenai sentimen analisis pada *tweet* novel yang dilakukan oleh Song [4] pada tahun 2017, juga menghasilkan hasil akurasi yang tinggi. Penelitian lain yang dilakukan oleh Alves [5] dengan studi kasus berupa *tweet* FIFA 2013 juga telah dilakukan, dengan membandingkan Naive Bayes dan Support Vector Machine (SVM). Selanjutnya, terdapat penelitian yang dilakukan oleh Guia [6] pada tahun 2019 dilakukan perbandingan berbagai classifier sentimen analisis seperti Naive Bayes, SVM, Decision Tree, dan Random Forest.

Penelitian yang dilakukan oleh Novendri [7] tahun 2020 membahas *multi-aspect* sentimen analisis dengan menggunakan 4 dataset berupa komentar YouTube memperoleh hasil akurasi yang tinggi. *Multi-aspect* sentimen analisis memungkinkan pengguna dan perusahaan untuk menentukan aspek mana yang perlu diperhatikan dari suatu produk. Seperti pada penelitian yang dilakukan oleh Ananda [8] pada tahun 2021, para peneliti menggunakan 5 aspek kategori dari ulasan hotel untuk membantu pengunjung menentukan hotel untuk mereka.

Penelitian yang dilakukan [9], [10] membahas tentang bagaimana membandingkan word embeddings termasuk Word2Vec dan GloVe pada analisis sentimen. Penelitian yang dilakukan oleh Shi dan Lie [11], membahas hubungan antara Word2Vec dan GloVe dalam klasifikasi teks dan karakteristik masing-masing metode tersebut. Dalam studi ini, sistem akan diklasifikasikan menggunakan Klasifikasi Naive Bayes karena klasifikasi ini memberikan hasil yang akurat dalam penelitian analisis sentimen [5], [6]. Kontribusi utama dari penelitian ini adalah membandingkan 2 ekstraksi fitur, yaitu: Word2Vec dan GloVe dalam kasus analisis sentimen multi-aspek yang belum dilakukan dalam penelitian sebelumnya. Dataset tidak hanya dibagi berdasarkan polaritas tetapi juga dibagi menjadi 3 aspek, yaitu: performa, harga, dan ketersediaan. Selain itu, beberapa algoritma Klasifikasi Naive Bayes juga akan digunakan yang akan dievaluasi menggunakan matriks kebingungan untuk melihat algoritma mana yang memberikan hasil terbaik pada setiap ekstraksi fitur.

### Topik dan Batasannya

Topik dan batasan dalam penelitian ini adalah melakukan perbandingan ekstraksi fitur Word2Vec dan GloVe menggunakan Naive Bayes Classifier pada *multi-aspect* sentimen analisis terhadap *tweet* produk Nvidia RTX dengan hashtag #nvidiartx. *Tweet* akan dikelompokkan secara manual menjadi tiga aspek kategori yaitu: performa, harga, dan ketersediaan karena saat ini belum terdapat alat bantu atau tools otomatis untuk pelabelan kategori yang efektif dan akurat. *Tweet* yang digunakan dibatasi dari rentang waktu 1 Januari 2020 hingga 25 Oktober 2022 sehingga jumlah data yang berhasil dikumpulkan sebanyak 1406 *tweet* yang akan diproses dalam dua skenario yaitu dengan tidak mengelompokkan data berdasarkan aspek kategori dan dengan mengelompokkan data berdasarkan aspek kategori yang telah ditentukan.

### Tujuan

Tujuan dari penelitian ini adalah untuk membandingkan performansi dari ekstraksi fitur Word2Vec dengan GloVe dalam melakukan *multi-aspect* sentimen analisis terhadap *tweet* produk Nvidia RTX dengan hashtag #nvidiartx dengan mengelompokkan *tweet* tersebut ke dalam tiga aspek kategori menggunakan Naive Bayes

Classifier.