

Multi-aspect Sentiment Analysis Terhadap Nvidia...

By: Igd Raditya Wibhawa Mn

As of: Nov 9, 2022 10:57:43 AM
4,001 words - 19 matches - 13 sources

Similarity Index
4%

Mode: Similarity Report ▼

paper text:

<p>Multi-Aspect Sentiment Analysis terhadap Nvidia RTX dengan Support Vector Machine dan Word Embeddings</p>	4
---	---

Abstrak—Beberapa bulan belakangan ini, tren meningkatnya pembelian GPU terutama seri RTX kian meningkat, banyak orang berlomba lomba untuk membuat PC untuk meningkatkan kinerja dengan kondisi WFH(Work From Home) selama terjadinya pandemi. Peningkatan terhadap penggunaan GPU seri RTX ini tentunya mengundang pendapat dari berbagai macam kalangan. Hal ini menciptakan pro dan kontra di kalangan PC enthusiast, terutama di beberapa platform Media Sosial. Dari munculnya pro dan kontra ini, dapat dilakukan ekstraksi emosi terhadap komentar yang mengandung pro dan kontra tersebut. Metode ini disebut juga dengan Analisis Sentimen, tahapan yang perlu dilakukan untuk melakukan Analisis Sentimen ini yang pertama adalah preprocessing yang terdiri dari Noise Removal, Tokenizing, Stopword Removal. Kemudian dilanjutkan dengan proses kedua yang kedua, yaitu Ekstraksi Fitur dengan Word Embeddings dan Support Vector Machine. Berdasarkan Analisis Performansi pada model yang telah dibuat, menunjukkan bahwa pada penelitian kali ini SVM kernel Linear menunjukkan hasil yang terbaik pada seluruh dataset, sedangkan kernel rbf pada penelitian ini menunjukkan hasil yang kurang memuaskan sekalipun telah melakukan perubahan parameter pada Gamma dan C pada penelitian ini. Penelitian ini dibuat bertujuan untuk menganalisa performansi model dalam mengklasifikasikan sentimen sebuah teks dengan model SVM. Kata kunci— Support Vector Machine, Preprocessing, Word Embeddings I. Pendahuluan Beberapa bulan belakangan ini, tren meningkatnya pembelian GPU terutama seri RTX kian meningkat, banyak orang berlomba lomba untuk membuat PC untuk meningkatkan kinerja dengan kondisi WFH(Work From Home) selama terjadinya pandemi. Peningkatan terhadap penggunaan GPU seri RTX ini tentunya mengundang pendapat dari berbagai macam kalangan. Hal ini menciptakan pro dan kontra di kalangan PC enthusiast, terutama di beberapa platform Media Sosial. Dari munculnya pro dan kontra ini, dapat dilakukan ekstraksi emosi terhadap komentar yang mengandung pro dan kontra tersebut. Metode ini disebut juga dengan Analisis Sentimen, tahapan yang perlu dilakukan untuk melakukan Analisis Sentimen ini yang pertama adalah preprocessing yang terdiri dari Noise Removal, Tokenizing, Stopword Removal. Kemudian dilanjutkan dengan proses kedua yang kedua, yaitu Ekstraksi Fitur dengan Word Embeddings. Terakhir menggunakan model Support Vector Machine terhadap hasil dari Ekstraksi Fitur Word Embeddings yang telah dihasilkan sebelumnya. Berdasarkan Analisis Performansi pada model yang telah dibuat, menunjukkan bahwa pada penelitian kali ini SVM kernel Linear menunjukkan hasil yang terbaik pada seluruh dataset, sedangkan kernel rbf pada penelitian ini menunjukkan hasil yang kurang memuaskan sekalipun telah melakukan perubahan parameter pada Gamma dan C pada penelitian ini. Bagaimana pemodelan dengan

Word Embeddings dan Support Vector Machine terutamanya kernel Linear mampu mengklasifikasikan sebuah teks dengan baik dan cukup akurat menunjukkan penelitian ini sesuai dengan yang diharapkan. Penelitian ini dibuat bertujuan untuk menganalisa performansi model dalam mengklasifikasikan sentimen sebuah teks dengan model SVM.

Berdasarkan latar belakang diatas dapat dirumuskan masalah antara lain

5

adalah berapa akurasi dan skor dari berbagai macam kernel pada svm terhadap klasifikasi data multi -aspect dan aspek mana yang memiliki akurasi SVM terbesar. Bertujuan untuk mengukur skor evaluasi dan akurasi masing masing kernel dan menemukan aspek dengan akurasi SVM terbesar. II. Kajian Teori A. Document Level Sentiment Classification Berikut adalah beberapa penelitian yang akan dicantumkan dibawah sebagai perbandingan dengan penelitian yang sekarang, berhubungan dengan Analisis Sentimen, namun menggunakan beberapa Text Preprocessing, Feature Extraction, dan pemodelan Machine Learning yang berbeda. Pada penelitian [2] menggunakan dataset yang telah ditentukan Sentimen Positif dan Negatif. Pemodelan

Machine Learning yang digunakan adalah Support Vector Machine dan Naïve Bayes . Text

1

Preprocessing **yang** digunakan **adalah**

N-Grams yaitu Bigrams. Hasil yang didapatkan menunjukkan accuracy dari Naïve Bayes memiliki skor yang lebih tinggi dibandingkan dengan Support Vector Machine. Pada penelitian [17] menggunakan dataset yang di crawl dari twitter berhubungan dengan angkutan publik dan telah ditentukan sentimen Positif dan Negatifnya. Pemodelan Machine Learning

yang digunakan adalah Support Vector Machine . Text Preprocessing **yang digunakan adalah**

6

POS(Part of Speech) Tagging dengan HMM dan Rule Based POS Tagging, kemudian di weighting dengan TF-IDF. Pada penelitian [1], [18], [11] menggunakan dataset yang diambil berdasarkan masing masing topik dimana telah ditentukan sentimen Positif dan Negatifnya. Penelitian [1] menggunakan Bengali dataset ABSA, Penelitian [18] menggunakan dataset berhubungan dengan Smartphone, penelitian [11] menggunakan dataset yang berhubungan dengan self- driving cars. Pemodelan

Machine Learning yang digunakan oleh ketiganya **adalah Support Vector Machine** . Text

8

Preprocessing **pada**

[1] menggunakan N- Grams, sedangkan pada [18] menggunakan Part-Of- Speech(POS) Tagging. Feature Extraction yang digunakan ketiganya adalah TF-IDF. Pada penelitian [19], [20], [21] menggunakan dataset masing masing topik, Sentimen Positif dan Negatif telah ditentukan. Penelitian [19] menggunakan dataset yang di crawl dari Twtter berkaitan dengan COVID-19, penelitian [20] menggunakan Cornell Movie Review Dataset, penelitian [21] menggunakan dataset review google play store pada aplikasi by.U. Pemodelan

Machine Learning yang digunakan [19] dan [21] **adalah Support Vector Machine**

1

, sedangkan penelitian [20] menggunakan K-Nearest Neighbor. Feature Extraction yang digunakan oleh ketiganya adalah TF-IDF. Pada penelitian [22] menggunakan Dataset yang di crawl dari twitter yang berhubungan dengan brimob dan mako brimob, sentimen Positif dan Negatif ditentukan dengan cara manual. Pemodelan

Machine Learning yang digunakan adalah Support Vector Machine dan Naïve Bayes . Text

1

Preprocessing **yang** digunakan **adalah**

tokenization, text normalization, dan stopword removal, tanpa menggunakan Feature Extraction manapun. Hasil Akurasi dan Classification Report Support Vector Machine sedikit lebih tinggi dibandingkan dengan Naïve Bayes. Pada penelitian [23] menggunakan dataset dari Kaggle yang berisi review mobile phone sebanyak 400rb data, sentimen Positif dan Negatif di labelling secara manual. Pemodelan

Machine Learning yang digunakan adalah Support Vector Machine, Naïve Bayes

1

, Decision Tree, Random Forest. Menggunakan TF-IDF setelah melakukan Text Preprocessing. Hasil yang paling tinggi didapatkan oleh Support Vector Machine dibandingkan dengan Classifier lainnya. B. Analisis Sentimen

Analisis Sentimen merupakan bidang studi yang menganalisa **tentang opini**

12

masyarakat, sentimen dan emosi terhadap suatu produk yang berkaitan terhadap suatu produk, isu sosial, acara, dll. Hal ini sangat penting karena dapat membantu untuk mengetahui apakah seseorang menyukai atau tidak menyukai suatu produk[1]. Tiap opini yang dikeluarkan oleh masyarakat melalui sosial media ini memiliki polaritas. Polaritas merujuk pada bentuk dasar, yang menunjukkan apakahteks tersebut positif atau negatif[2].Penelitian terhadap penggunaan Twitter menunjukkan bahwa19% dari tweets merujuk pada sebuah brand atau produk, 20% tweet juga menunjukkan sebuah ekspresi sentiment dari suatu brand[12]. Namun teks dari sosial media ini tidak terstruktur denganbaik sehingga menjadi

tantangan untuk Analisis Sentimen. Teks tersebut biasanya bermasalah itu panjang teks, misspelling, penggunaan slang, dan kata yang diperpendek[9]. C. Aspect Based Sentiment Analysis

Aspect Based Sentiment Analysis merupakan salah satu Analisis Sentimen yang

7

mendeteksi sebuah sentimen terhadap aspek tertentu dalam sebuah kata. Kalimat yang memiliki beberapa aspek ini dapat dikategorikan sebagai Overlapping atau Non- Overlapping. Sebuah kalimat dapat dikatakan sebagai Non-Overlapping apabila dua dari aspek kalimat tersebut tidak tumpang tindih[7]. Aspect Based Sentiment Analysis bertujuan untuk mengenali polaritas dari sebuah sentiment terhadap aspek spesifik dari sebuah kalimat. Target dari aspek merujuk pada sebuah kata atau frasa yang mendeskripsikan aspek dari sebuah entitas[8]. D.

Support Vector Machine Support Vector Machine merupakan salah satu model Machine Learning
untuk Klasifikasi **yang**

1

biasanya sering digunakan pada Analisis Sentimen. Kemampuan dari SVM itu sendiri adalah kemampuan untuk membaca pola data klasifikasi dengan akurasi dan reproduktifitas yang tinggi. Decision Function dari SVM adalah memilih "Hyperplane" yang paling optimal yang berfungsi untuk memisahkan observasi yang dimiliki oleh satu kelas dengan kelas lainnya berdasarkan pola informasi dari observasi tersebut yang disebut feature[3]. SVM telah digunakan untuk masalah two- class classification dan dapat diaplikasikan pada kernel linear dan non-linear. Non- Linear (polynomial, radial basis, sigmoid) classifier menggunakan berbagai macam kernel untuk memaksimalkan margin antara hyperplane[13]. Beberapa kekurangan dari Support Vector Machine antara lain adalah biaya komputasi pada data set yang besar karena training kernel matrix tumbuh dengan bentuk quadratic dari ukuran dataset, dan tidak mudah untuk menyelesaikan masalah dengan multi- classification problem[10]. E. Word Embeddings Word Embeddings merupakan salah satu metode NLP (Natural Language Processing) yang sering digunakan untuk mengolah teks, biasanya digunakan pada Sentimen Analisis, Recommender System, dll. Word Embeddings adalah sebuah teknik untuk merepresentasikan sebuah teks menjadi sebuah vektor numerik yang bertindak sebagai vektor dari sebuah kata[4]. Word Embeddings memiliki 2 kategori untuk metode evaluasi yaitu adalah Extrinsic Evaluator yang menggunakan Word Embeddings sebagai Input, Analisis Sentimen adalah salah satu contohnya, yang satunya adalah Intrinsic Evaluator yang mengukur hubungan semantik antara sebuah kata secara langsung[5]. Tidak hanya digunakan untuk Classification, namun Word Embeddings juga dapat digunakan sebagai untuk Clustering dan Regression[6]. III.

Metode A. Flowchart Sistem Gambar 1 - Flowchart Sistem

13

Pada flowchart diatas, akan dilakukan beberapa tahap dalam penelitian ini, yang pertama adalah penentuan aspek dari produk yang akan diteliti sentimennya, setelah menentukan aspek kemudian melakukan scraping data, setelah di scraping, akan ditentukan label dari teks yang telah di scrap apakah termasuk ke dalam sentiment positif, negatif, atau netral. Setelah itu

akan melalui tahap Pre-processing yang terdiri dari

10

Lowercase and Noise Removal, Tokenization, Stopword Removal, dan Word Embeddings. Kemudian membuat model yang mengklasifikasikan sentiment sebuah teks. Terakhir mengukur hasil Evaluasi dari keempat kernel yang digunakan dari SVM. Detail tiap tahap dapat dilihat pada poin poin dibawah. B. Penentuan Aspek Sebelum melakukan Scraping Data, perlu melakukan analisa terlebih dahulu apa saja yang menjadi faktor/aspek penting ketika seorang Customer memutuskan untuk membeli GPU Nvidia seri RTX. RTX sendiri adalah salah satu generasi baru GPU keluaran Nvidia pada tahun 2020. RTX sendiri baik dari segi fitur jauh lebih baik dibanding pendahulunya, dimana GPU ini memiliki performa yang sangat kuat, memiliki fitur terbaru yang disebut dengan DLSS untuk meningkatkan performa ketika gaming, dan Ray Tracing yang mampu meningkatkan kualitas visual sehingga terlihat nyata. Berdasarkan hal ini, akan di scraping sentimen yang berhubungan dengan aspek fitur Ray-Tracing, aspek fitur DLSS, dan aspek fitur Performa. Penentuan teks itu sendiri apakah termasuk ke dalam aspek Ray Tracing/DLSS/Performa, tergantung dari data yang di scrap, apabila saat melakukan scraping yang berkaitan dengan harga kemudian teks/sentiment tersebut muncul tanpa harus menyinggung terkait harga secara langsung, maka akan termasuk kedalam dataset harga, begitu pula dengan aspek yang lainnya. Contohnya adalah sebagai berikut : Gambar 2 - Contoh Video Gambar 3 - Contoh Comment Gambar 2 adalah sebuah video yang memperkenalkan fitur DLSS. Gambar kedua adalah salah satu komentar yang ada pada video tersebut. Komentar tersebut akan masuk kedalam dataset fitur DLSS karena saat scraping komen tersebut akan muncul, dan secara tidak langsung, komentar tersebut menyinggung tentang fitur DLSS. Gambar 4 - Pie Chart Aspek Ray-Tracing Gambar 5 - Pie Chart Aspek Performa kedalam sentimen positif, negatif, atau netral. Contohnya bisa dilihat seperti tabel dibawah : Tabel 1 - Text Sentiment Teks Sentimen nobody is going to talk Netral(3) about the thumbnail Gambar 6 - Pie Chart Aspek DLSS sadly games are more

Pada Gambar 4, Gambar 5, dan Gambar 6

3

, beautiful than ever but Negatif(2) menunjukkan komposisi persentasi jumlah data positif, theyre getting shittier as negative, dan netral di masing masing aspek. well for people like me who Positif(1) C. Scraping Data are used to play since Pada bagian ini akan melakukan scraping data yang kid in low resolutions diperlukan dari berbagai sumber. Pada step ini, sumber and low quality dlss is yang akan diambil datanya berasal dari Twitter dan like a wonder Youtube, dengan menggunakan API yang disediakan Jumlah sentiment pada masing masing aspek oleh website Netlytic.org untuk mengumpulkan sebagai berikut : sentiment yang diperlukan. Berikut adalah tampilan dari website Netlytic.org yang akan digunakan untuk Tabel 2 - Jumlah Data per Aspek melakukan scraping data : Aspek Jumlah Jumlah Jumlah Jumlah Data

Positif Negatif Netral Fitur 2305 782 754 data 769 data Ray data data Tracin g Fitur 2184 729 715 data 740 data Perfor data data Gambar 7 - Netlytic Youtube API ma Gambar 7 adalah bagian dari Netlytic.org yang Fitur 2545 852 854 data 839 data menggunakan youtube API untuk scraping comment DLSS data data dari suatu video. E. Text Preprocessing Lowercase dan Noise Removal Pada bagian ini akan menghilangkan simbol simbol, serta mengubah semua huruf menjadi kecil. Berikut adalah daftar simbol yang dihapus dari sebuah teks: Gambar 8 - Netlytic Twitter API Gambar 8 adalah bagian dari Netlytic.org yang Tabel 3 - Punctuation menggunakan Twitter API untuk scraping tweet dari seseorang. ! () - = ? { \ # ' * : > @ } ^ D. Labelling " & + / < [| _ Pada tahap ini akan melakukan labelling secara manual dengan menentukan apakah teks tersebut masuk \$ % , ; ; ~ ` [oh,my,god,it,is,so,good] [oh,god,good] Tokenization Kemudian melakukan Tokenizing, pada bagian ini memerlukan akses tiap huruf dalam string. Maka kita Word Embeddings memerlukan Tokenization untuk memecah teks didalamnya menjadi komponen yang lebih kecil, Yang terakhir adalah pengaplikasian Word diperlukan untuk mencari jumlah kata dan kalimat Embeddings. Word Embeddings sendiri adalah teknik didalam text, menentukan berapa kali kata yang spesifik yang mengubah suatu teks menjadi sebuah vektor. atau frasa tertentu muncul. Vektor tersebut merepresentasikan sebuah kalimat yang telah diubah sebelumnya. Word Embeddings yang Berikut adalah salah satu contoh teks yang telah digunakan adalah spaCy.vector yang menggunakan melewati proses Tokenizing : algoritma dari GloVe untuk mengubah suatu kalimat Tabel 4 - Contoh Tokenization menjadi sebuah vector. Berdasarkan dari [25], pre-trained Word Text Tokenized Text Embeddings memiliki bentuk 300 dimensi, dimana keseluruhan kata didalam teks ini merepresentasikan if you want to get a good [im,glad,im,not,looking,f 300 dimensi yang ada, contoh input dan output teks graphics card with good or,a,gpu,right, yang telah mengalami vektorisasi sebagai berikut : price just buy it from now,mincraft,runs,at,ab vietnam you can get a out,smoothly, radeon gb for just bucks with,high,settings,so,im, good,rn] im glad im not looking [im,glad,im,not,looking,f for a gpu right now or,a,gpu,right, minecraft runs at about now,mincraft,runs,at,ab smoothly with high out,smoothly, settings so im good rn with,high,settings,so,im, good,rn] Gambar 9 - Input dan Output Word Embeddings F. Classification Model with SVM oh my god it is so good [oh,my,god,it,is,so,good] Pada tahap ini akan membuat model klasifikasi menggunakan Support Vector Machine. Support Vector Machine sendiri bekerja dengan cara menganalisis data Stopword Removal dan menentukan decision boundaries dengan memiliki hyper-plane, yang mana hyper-plane ini memisahkan Kemudian menggunakan Stopword Removal, vektor dari satu kelas dengan kelas lain[24]. SVM bertujuan untuk menghilangkan sebuah kata yang tidak menyediakan hasil akurasi yang lebih baik dalam memiliki arti atau informasi penting didalamnya, kamus mengklasifikasikan teks karena SVM sendiri termasuk stopwords yang digunakan adalah kamus English tanpa binary classifier yang sering digunakan dalam Text memodifikasi/menambah kata yang perlu dihilangkan. Classification sehingga cocok digunakan sebagai untuk mengklasifikasikan polaritas dari sebuah kalimat[1]. Berikut adalah contoh salah satu teks yang telah Walaupun SVM adalah sebuah binary classifier, hasil melalui stopwords removal: yang diharapkan akan tinggi sekalipun Classifiernya lebih dari dua, untuk kernel yang digunakan nantinya Tabel 5 - Contoh Stopwords akan menggunakan kernel yang memiliki akurasi yang tinggi seperti yang ditunjukkan nanti pada bagian Evaluation. Ketika melakukan fit dataset, yang akan Tokenized Stopword Removal menjadi X adalah teks yang telah di cleaning dan di vektorisasi, Y adalah nilai sentimennya. [yeah,running,ti,to,and,it [yeah,running,ti,super,go s,super,good,i od,idk,people,d Berikut adalah code yang digunakan dalam dk,why,people,just,dont, ont,used,ti] membuat model Klasifikasi dengan SVM untuk by,used,ti] Analisis Sentimen : [im,glad,im,not,looking,f [im,glad,im,looking,gpu, or,a,gpu,right, right,mincraft,r now,mincraft,runs,at,ab uns,smoothly,high,settin out,smoothly,

gs,im,good,rn] Gambar 10 - Train Test Split dan Data Train with,high,settings,so,im, Pada Gambar 10, dilakukan terlebih dahulu good,rn] train_test_split untuk menentukan data train dan data test, untuk parameter yang diberikan data train berukuran 80% dari keseluruhan data, sedangkan data laptop test sebanyak 20%, kemudian random_state = 6 menunjukkan bahwa data akan di shuffle sebanyak 6 and still using kali. Kemudian SVM yang digunakan adalah kernel core i without Linear. graphics card c crisis is so bad 2 2 Tabel 6 - Contoh Hasil Prediksi dan Aktual it hasnt a walk animation Teks Hasil Prediksi Aktual my pc crying 1 2 3 3 after he seen it probably this video took a big that was so 1 1 effort to make fuckin good this not by that i just start you by your sleeping so pc rip relax nice but 1 1 your pc nothing against ray tracing minecraft has 1 2 its own feeling realistic Tabel 6 menunjukkan beberapa teks yang graphic looks dibandingkan antara hasil prediksi, dengan nilai yang good but it aktual/sebenarnya doesnt give us feeling of minecraft that G. Evaluation blocky game Pada bagian ini akan menampilkan Classification Report

yang terdiri dari Precision, Recall, dan F1-Score

9

, i cant even get 1 1 dimana masing masing metrik tersebut memiliki rumus fps normally sebagai berikut: without shaders lol 1. Precision TP whats rtx 3 3 ($TP+FP$) (1) 3 1 2. Recall im watching TP this video in p ($TP+FN$) (2) because my 3. F1-Score computer cannot stand $2 \times \text{Precirinn} \times \text{Recall}$ $\text{Precirinn} + \text{Recall}$ (3) even Pada bagian Evaluation ini, akan menggunakan Cross Validation terhadap 4 kernel yang akan dipilih this video lmao sebagai bagian dari SVM, yaitu kernel rbf, Sigmoid, iwtd Polynomial, serta Linear. Berikut akan menampilkan 3 3 hasil klasifikasi dari salah satu dataset, yaitu dataset everyone just DLSS : had epic pc Hasil Klasifikasi Aspek fitur Ray-Tracing : and im just Tabel 7 - Hasil Klasifikasi Aspek fitur Ray-Tracing sitting here with my old Positif Negatif Netral RBF 437 1 23 Sigmoid 140 131 191 Polynomial 147 133 181 Linear 187 147 127 Hasil Klasifikasi Aspek fitur Performa : Tabel 8 - Hasil Klasifikasi Aspek fitur Performa Positif Negatif Netral RBF 397 6 34 Sigmoid 131 101 205 Polynomial 105 152 180 Linear 170 132 135 Hasil Klasifikasi Aspek fitur DLSS : Tabel 9 - Hasil Klasifikasi Aspek fitur DLSS Positif Negatif Netral RBF 2 497 10 Sigmoid 143 130 236 Polynomial 143 179 187 Linear 204 151 154 IV. Hasil dan Pembahasan Telah dilakukan eksperimen terhadap 3 macam dataset dan pemodelan Machine Learning yang telah dipilih, yaitu Support Vector Machine. 3 Dataset ini terdiri merupakan faktor yang berbeda beda, yang pertama adalah faktor Performa, kedua adalah faktor Ray Tracing, dan ketiga adalah faktor DLSS. Berikut akan ditampilkan

precision, recall, dan f1-score dan akurasi dari masing masing

2

dataset A. Dataset Performance Pada dataset Scalper ini terdapat sebanyak 2997 data, setting gamma = 1 dan c = 1(default), cv = 10 ,kemudian untuk akurasinya terdapat 4 akurasi, yang pertama menggunakan kernel rbf dengan akurasi 41%, kemudian menggunakan kernel sigmoid dengan akurasi 51%, kemudian dengan kernel polynomial dengan akurasi 61%, dan kernel linear dengan akurasi 64%. Berikut adalah jumlah masing masing sentimen,

precision, recall, dan f1-score dari masing masing

2

kernel. Jumlah Data : Tabel 10 - Jumlah Dataset Performa Label Jumlah 1(Positive) 729 2(Negative) 715 3(Neutral) 740 Total 2184 Akurasi dan Metrik : Tabel 11 - Akurasi dan Metrik Dataset Performa Kernel Metrik Recall F1- Score Akurasi RBF(Gamma = 1, C = 1) 0.36 0.38 0.91 0.90 0.11 0.20 0.52 0.18 0.33 0.41 = 41% Sigmoid 0.55 0.60 0.44 0.52 0.40 0.60 0.53 0.48 0.51 0.51 = 51% Polynomial 0.70 0.60 0.56 0.54 0.54 0.74 0.61 0.57 0.64 0.61 = 61% Linear 0.62 0.60 0.72 0.72 0.56 0.64 0.67 0.58 0.68 0.64 = 64% Pada parameter Gamma dan C yang dimana hanya digunakan oleh kernel RBF pada SVM, menunjukkan bahwa pada perubahan nilai Gamma dan C mulai dari Gamma = 1, C = 1 hingga Gamma = 100, C = 100, tidak terjadi perubahan skor akurasi seperti pada grafik berikut: Gambar 11 - Hasil Perubahan skor Gamma dan C pada RBF Gambar 12 - Visualisasi Evaluasi Dataset Performa(Average) Di bagian ini kernel Linear memiliki metrik paling tinggi secara keseluruhan, dari keseluruhan metrik, skor precision paling tinggi, persebaran skor pada Sigmoid, Polynomial, dan Linear seimbang. Pada RBF, Skor recall dan F1-Score cukup jauh dibanding Precision. B. Dataset Ray Tracing Pada dataset harga terdapat sebanyak 2100 data, setting gamma = 1 dan c = 1(default), cv = 10 ,terdapat 4 akurasi, yang pertama adalah kernel rbf dengan akurasi 40%, kemudian kernel sigmoid dengan akurasi 51%, kernel polynomial dengan akurasi 62%, kernel linear dengan akurasi 63%. Berikut adalah jumlah masing masing sentimen, precision, recall, dan f1 dari masing masing kernel. Jumlah Data : Tabel 12 - Jumlah Dataset Ray-Tracing Label Jumlah 1(Positive) 782 2(Negative) 754 3(Neutral) 769 Total 2305 Akurasi dan Metrik : Tabel 13 - Akurasi dan Metrik Dataset Ray-Tracing Kernel Metrik Akurasi Precision Recall F1- Score RBF(Gamma = 1, C = 1) 0.36 0.67 1.00 0.01 0.53 0.01 0.36 = 36% 0.95 0.17 0.29 Sigmoid 0.58 0.51 0.54 0.44 0.56 0.47 0.46 = 46% 0.47 0.57 0.51 Polinomial 0.67 0.59 0.59 0.53 0.63 0.56 0.63 = 63% 0.61 0.74 0.67 Linear 0.61 0.59 0.67 0.59 0.64 0.59 0.65 = 65% 0.71 0.63 0.67 Pada parameter Gamma dan C, yang hanya digunakan pada kernel RBF, menunjukkan bahwa perubahan nilai Gamma dan C, mulai dari Gamma = 1, C = 1, hingga Gamma = 100, C = 100, tidak memiliki pengaruh terhadap perubahan skor akurasi seperti pada grafik berikut : Gambar 13 - Hasil Perubahan skor Gamma dan C pada RBF Gambar 14 - Visualisasi Evaluasi Dataset Ray-Tracing(Average) Di bagian ini, kernel Linear memiliki skor Precision, Recall, F1-Score paling tinggi dibandingkan kernel lainnya. Persebaran skor evaluasi dari kernel Sigmoid, Polynomial, dan Linear terlihat seimbang, terkecuali kernel RBF, dimana skor Recall dan F1- Score berbeda jauh dibandingkan dengan skor Precision. C. Dataset DLSS Pada dataset ini terdapat sebanyak 2400 data, setting gamma = 1 dan c = 1(default), cv = 10,terdapat 3 akurasi dari masing masing kernel, yang pertama dengan kernel rbf dengan akurasi 36%, kemudian dengan kernel sigmoid dengan akurasi 46%, kernel polynomial dengan akurasi 63%, kernel linear dengan akurasi 65%. Berikut adalah jumlah masing masing sentimen,

precision, recall, dan f1- score dari masing masing

2

kernel. Jumlah Data : Tabel 14 - Jumlah Dataset DLSS Label Jumlah 1(Positive) 852 2(Negative) 854 3(Neutral) 839 Total 2545 Akurasi dan Metrik : Tabel 15 - Akurasi dan Metrik Dataset DLSS Kernel Metrik Akurasi Precision Recall F1- Score RBF(Gamma = 1, C = 1) 0.35 0.34 0.40 0.60 0.37 0.44 0.36 = 36% 0.97 0.09 0.16 Sigmoid 0.55 0.46 0.50 0.46 = 46% 0.50

0.42 0.46 0.39 0.51 0.44 Polinom ial 0.72 0.60 0.56 0.58 0.63 0.59 0.63 = 63% 0.60 0.75 0.67 Linear 0.64 0.72 0.68 0.65 = 65% 0.62 0.60 0.61 0.70 0.63 0.66 Pada parameter Gamma dan C, yang hanya digunakan pada kernel RBF, menunjukkan bahwa perubahan nilai Gamma dan C, mulai dari Gamma = 1, C = 1, hingga Gamma = 100, C = 100, tidak memiliki pengaruh terhadap perubahan skor akurasi seperti pada grafik berikut: Gambar 15 - Hasil Perubahan skor Gamma dan C pada RBF Gambar 16 - Visualisasi Evaluasi Dataset DLSS(Average) Di bagian ini, kernel Linear memiliki skor Precision, Recall, F1-Score paling tinggi dibandingkan kernel lainnya. Persebaran skor evaluasi dari kernel Sigmoid, Polynomial, dan Linear terlihat seimbang, terkecuali kernel RBF, dimana skor Recall dan F1- Score berbeda jauh dibandingkan dengan skor Precision. Berikut adalah visualisasi akurasi dari 3 dataset/aspek yang menggunakan 4 macam kernel : Gambar 17 - Visualisasi Akurasi 3 Dataset Berdasarkan hasil yang telah ditunjukkan pada masing masing kernel, kernel Linear unggul pada semua Dataset. Dibawah kernel Linear ada kernel Polynomial yang akurasinya tidak berbeda jauh dengan kernel Linear. Sedangkan kernel RBF dan kernel Sigmoid, akurasi yang dihasilkan kurang memuaskan, parameter yang ditambahkan pada kernel RBF dimulai dengan Gamma = 1 C = 1, hingga Gamma = 100 C = 100 tidak memberikan hasil yang signifikan. Diantara 3 aspek, yang memiliki akurasi tertinggi dari keempat kernel pada gambar diatas adalah aspek DLSS dengan menggunakan akurasi hingga 65% dengan kernel Linear. Dengan membandingkan penelitian lain seperti [26] dan [27] menunjukkan bahwa kernel Linear unggul pada Analisis Sentimen dibandingkan dengan kernel lainnya. Namun pada penelitian [28] menunjukkan hal berbeda, dimana yang paling unggul adalah kernel RBF. Pada grafik yang membandingkan skor dari kernel menunjukkan bahwa, kernel RBF memiliki skor Recall dan F1-Score yang cukup rendah dibandingkan dengan Precision. Skor Recall yang rendah menunjukkan bahwa teks dengan sentiment positif yang diprediksi positif tidak banyak. Menunjukkan bahwa Recall rendah dapat mempengaruhi F1-Score. V. Kesimpulan Pada penelitian kali ini yang membahas tentang Multi-Aspect Sentiment Analysis terhadap NVIDIA RTX Series telah menunjukkan beberapa hasil penerapan model Machine Learning Support Vector Machine di beberapa kernel. Untuk pembahasan Analisis Sentimen terhadap Produk Nvidia RTX kali ini menunjukkan bahwa SVM dengan kernel Linear menunjukkan hasil terbaik di keseluruhan dataset dengan akurasi sebanyak 64% pada dataset Performance, 63% pada dataset Ray Tracing, dan 65% pada dataset DLSS. Melakukan perbandingan dengan paper lain juga menunjukkan kernel Linear sering menunjukkan hasil terbaik dalam Analisis Sentimen, namun ada pula selain kernel Linear, contohnya adalah kernel RBF yang menunjukkan hasil yang terbaik pada penelitian

lain. Hal ini dipengaruhi beberapa faktor, diantaranya adalah

11

Text Preprocessing, Feature Extraction, dataset yang digunakan, dsb. Penelitian Multi-Aspect Sentiment Analysis terhadap Nvidia RTX dapat dikembangkan lebih jauh lagi dengan seri GPU yang lain, serta menggunakan model dan metode yang berbeda juga.

sources:

1

56 words / 1% - Internet from 13-Aug-2018 12:00AM
id.123dok.com

- 2 24 words / 1% - Internet from 05-Jul-2019 12:00AM
www.coursehero.com

- 3 8 words / < 1% match - Internet from 11-Oct-2020 12:00AM
id.123dok.com

- 4 14 words / < 1% match - Internet from 13-Sep-2022 12:00AM
openlibrary.telkomuniversity.ac.id

- 5 9 words / < 1% match - Internet from 06-Jun-2019 12:00AM
bhaskaranet.blogspot.com

- 6 9 words / < 1% match - Internet from 04-Sep-2021 12:00AM
digilib.unhas.ac.id

- 7 9 words / < 1% match - Internet from 23-Oct-2021 12:00AM
ecampus.pelitabangsa.ac.id

- 8 9 words / < 1% match - Internet from 24-Aug-2020 12:00AM
repositori.usu.ac.id

- 9 8 words / < 1% match - Internet from 06-Jun-2022 12:00AM
github.com

- 10 8 words / < 1% match - Internet from 12-Jul-2018 12:00AM
media.neliti.com

- 11 8 words / < 1% match - Internet from 25-May-2021 12:00AM
ojs.uho.ac.id

- 12 8 words / < 1% match - Internet from 04-Aug-2021 12:00AM
openlibrarypublications.telkomuniversity.ac.id

- 13 8 words / < 1% match - Internet from 16-Oct-2020 12:00AM
simki.unpkediri.ac.id