



Published: 2021-04-25

Articles

Modifikasi Skema Teknik Tanam Padi dan Bajak Sawah Berbasis Square Transposition 64-bit
(<https://journal.maranatha.edu/index.php/jutisi/article/view/2973>)

Nadya Glorya Najoan, Magdalena Ariance Irene Pakereng

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/2973/1828>)

Sistem Rekomendasi Suku Cadang Berdasarkan Item Based Filtering
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3036>)

Christian Wibisono, Lucky Surya Haryadi, Juan Elisha Widyaya, Swat Lie Liliawati

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3036/1829>)

Continuous Integration and Continuous Delivery Platform Development of Software Engineering and Software Project Management in Higher Education
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3254>)

Sendy Ferdian, Tjatur Kandaga, Andreas Widjaja, Hapnes Toba, Ronaldo Joshua, Julio Narabel

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3254/1830>)

Perbandingan Algoritma Machine Learning dalam Menilai Sebuah Lokasi Toko Ritel
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3182>)

Kristiawan Kristiawan, Andreas Widjaja

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3182/1831>)

Prediksi Penyebaran Informasi di Twitter dengan Metode Pembelajaran Mesin dengan Fitur Linimasa
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3324>)

Lucky Surya Haryadi, Bernard Renaldy Suteja

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3324/1832>)

Form Recognition dan Character Mapping Menggunakan Image Segmentation dan Optical Character Recognition
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3340>)

Christian Wibisono, Setia Budi

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3340/1833>)

Pengembangan Knowledge Management System dengan Teknik Information Retrieval
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3316>)

Try Atmaja Linggan Jaya, Mewati Ayub

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3316/1834>)

Pendeteksian Citra Pengunjung Menggunakan Single Shot Detector untuk Analisis dan Prediksi Seasonality
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3329>)

Deon Diamanta, Hapnes Toba

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3329/1835>)

Pengaruh Preprocessing Terhadap Klasifikasi Diabetic Retinopathy dengan Pendekatan Transfer Learning Convolutional Neural Network
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3327>)

Juan Elisha Widyaya, Setia Budi

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3327/1836>)

Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3216>)

Ni Luh Putu Chandra Savitri, Rada Amirur Rahman, Reyhan Venyutzky, Nur Aini Rakhmawati

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3216/1837>)

Pengujian dan Model Penerimaan Aplikasi Web Mapping System Lahan Pertanian
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3336>)

Delfia Balisa, Rosa Delima, Antonius Rachmat Chrismanto, Halim Budi Santoso

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3336/1838>)

Perancangan Sistem E-Reporting Menggunakan ReactJS dan Firebase
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3098>)

Jeremy Panjaitan, Andrew Fernando Pakpahan

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3098/1839>)

Evaluasi Tata Kelola Guna Meningkatkan Kinerja Manajemen Teknologi Informasi Menggunakan Framework COBIT 5
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3260>)

Fitria Febriani, Augie David Manuputty

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3260/1840>)

Pembangunan Dashboard untuk Mendukung Analisis Kartu Rencana Studi dan Kartu Hasil Studi Mahasiswa
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3355>)

Debora Syebat Nazir, Gloria Virginia, Restyandito Restyandito, Agata Filiana, Andhika Galuh Prabawati

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3355/1841>)

Penerapan Metode Profile Matching untuk Pendukung Keputusan Pemilihan Manajer Information Technology
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3393>)

Anita Diana, Dwi Achadiani, Hendri Irawan

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3393/1842>)

Klasifikasi Lukisan Karya Van Gogh Menggunakan Convolutional Neural Network-Support Vector Machine
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3399>)

Yohannes Yohannes, Daniel Udjulawa, Febbiola Febbiola

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3399/1843>)

Manual Load Balancing pada Redundancy Link Menggunakan Multi-Group Hot Standby Router Protocol
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3403>)

Fajar Hariadi

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3403/1844>)

Aplikasi Edukasi Pemecahan Masalah pada Vulnerability Server Berbasis Web
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3405>)

Ahmad Syaichul Hadi, Mochamad Alfian Rosid

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3405/1845>)

Risk Management Analysis Using COBIT 4.1 at Vehicle Testing Management Information System
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3419>)

Resad Setyadi, Septian Anggoro

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3419/1846>)

Pengendalian Kualitas Air untuk Tanaman Hidroponik Menggunakan Raspberry Pi dan Arduino Uno
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3421>)

Ryan Ganesha Calibra, Irfan Ardiansah, Nurpilihan Bafdal

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3421/1847>)

The Implementasi Metode On-Page Search Engine Optimization untuk Meningkatkan Peringkat Website sebagai Hasil Pencarian Google (<https://journal.maranatha.edu/index.php/jutisi/article/view/3428>)

Antonius Antonius, Bernard Renaldy Suteja

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3428/1848>)

Pengembangan Sistem Informasi Asosiasi Jasa Konstruksi dengan Menerapkan Tanda Tangan Digital
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3431>)

Taufan Abdurrachman, Bernard Renaldy Suteja

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3431/1849>)

Manajemen Risiko Divisi Sistem Informasi Perguruan Tinggi Dengan Framework COBIT 5
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3434>)

Francis Anderson Kojongian, Mewati Ayub

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3434/1850>)

Penentuan Aspek yang Berpengaruh Terhadap Produk Smartphone Berdasarkan Ulasan Berbasis Teksual
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3466>)

Hetthroh Sagala, Hapnes Toba

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3466/1851>)

Identifikasi Risiko Program Maintenance dalam Pengelolaan Proyek Berbasis Agile Menggunakan Pohon Klasifikasi
(<https://journal.maranatha.edu/index.php/jutisi/article/view/3545>)

Billyanto Hendrik, Bernard Renaldy Suteja

PDF (<https://journal.maranatha.edu/index.php/jutisi/article/view/3545/1852>)

SERTIFIKAT

Direktorat Jenderal Pendidikan Tinggi, Riset dan Teknologi
Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi Republik Indonesia





Kutipan dari Keputusan Direktorat Jenderal Pendidikan Tinggi, Riset dan Teknologi
Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia

Nomor 225/E/KPT/2022
Peringkat Akreditasi Jurnal Ilmiah periode III Tahun 2022

Nama Jurnal Ilmiah
JuTISI (Jurnal Teknik Informatika dan Sistem Informasi)
E-ISSN: 24432229
Penerbit: Maranatha University Press

Ditetapkan Sebagai Jurnal Ilmiah

TERAKREDITASI PERINGKAT 3

Akreditasi Berlaku selama 5 (lima) Tahun, yaitu
Volume 8 Nomor 1 Tahun 2022 sampai Volume 12 Nomor 2 Tahun 2026
Jakarta, 07 December 2022
Ptt. Direktur Jenderal Pendidikan Tinggi,
Riset, dan Teknologi



Prof. Ir. Nizam, M.Sc., DIC, Ph.D., IPU, ASEAN Eng



(https://maranathaedu-my.sharepoint.com/:i:/g/personal/jutisi_it_maranatha_edu/EVfVK2kulh1Doui8_Xgt5I4BHQYD3TnJGelcrRJ0pdrAyw?e=mSP021)

SINTA-3 Certificate:
Please click the above image to see the original size.

Language

English (https://journal.maranatha.edu/index.php/jutisi/user/setLocale/en_US?source=%2Findex.php%2Fjutisi%2Fissue%2Fview%2F208)

Bahasa Indonesia (https://journal.maranatha.edu/index.php/jutisi/user/setLocale/id_ID?source=%2Findex.php%2Fjutisi%2Fissue%2Fview%2F208)

Menu

Short History (<https://journal.maranatha.edu/index.php/jutisi/sejarah>)

Focus & Scope (<https://journal.maranatha.edu/index.php/jutisi/focusandscope>)

Publishing Frequency (<https://journal.maranatha.edu/index.php/jutisi/fp>)

Editorial Policy (https://journal.maranatha.edu/index.php/jutisi/editorial_policy)

Open Access Policy (<https://journal.maranatha.edu/index.php/jutisi/kat>)

Review Policy (https://journal.maranatha.edu/index.php/jutisi/review_policy)

Digital Archiving & Preservation Policy (<https://journal.maranatha.edu/index.php/jutisi/kebijakanArsip>)

Publication Ethics (https://journal.maranatha.edu/index.php/jutisi/etika_pub)



Crossref (<https://search.crossref.org?q=2443-2229>)



OCLC WorldCat[®] ([https://www.worldcat.org/search?qt=worldcat_org_all&q=2443-](https://www.worldcat.org/search?qt=worldcat_org_all&q=2443-2210)

[2210](https://www.base-search.net/Search/Results?lookfor=2443-2210&type=all&page=1&l=en&oaboost=1&refid=dcpagene) (<https://www.base-search.net/Search/Results?lookfor=2443-2210&type=all&page=1&l=en&oaboost=1&refid=dcpagene>)



(<https://sinta.kemdikbud.go.id/journals/profile/3410>)

Template



([https://maranathaedu-](https://maranathaedu-my.sharepoint.com/:w/g/personal/jutisi_it_maranatha_edu/ETkogMrLJ01AgfAj3G6Yb_sBBsOcpjr3GLR_jQnrtdN9OA?e=yCrN9C)

[my.sharepoint.com/:w/g/personal/jutisi_it_maranatha_edu/ETkogMrLJ01AgfAj3G6Yb_sBBsOcpjr3GLR_jQnrtdN9OA?e=yCrN9C](https://maranathaedu-my.sharepoint.com/:w/g/personal/jutisi_it_maranatha_edu/ETkogMrLJ01AgfAj3G6Yb_sBBsOcpjr3GLR_jQnrtdN9OA?e=yCrN9C))

Tools



MENDELEY (<https://www.mendeley.com>)

(<https://www.grammarly.com>)

(<https://www.turnitin.com>)

Penentuan Aspek yang Berpengaruh Terhadap Produk *Smartphone* Berdasarkan Ulasan Berbasis Tekstual

<http://dx.doi.org/10.28932/jutisi.v7i1.3466>

Riwayat Artikel

Received: 26 Februari 2021 | Final Revision: 22 Maret 2021 | Accepted: 25 Maret 2021

Hetthroh Sagala^{#1}, Hapnes Toba^{✉*2}

[#]Program Studi Magister Ilmu Komputer, Fakultas Teknologi Informasi

Universitas Kristen Maranatha

Jl. Prof. drg. Surya Sumantri No.65 Bandung

¹1879014@maranatha.ac.id

²hapnestoba@it.maranatha.edu

Abstract— With the rapid development of new technologies in smartphones, understanding market trends has become an increasingly difficult task. In these circumstances, online product reviews that can reflect consumer sentiment about the product has been a concern for now. Online review analysis can help sellers understand consumer interests and desires prior to launching a new product. This research contributes to determine the state of the smartphone's market by creating a method to find important aspects of a smartphone. The data source comes from the Amazon e-commerce web with four predefined smartphone brands. In this study, the authors used topic modeling with the LDA algorithm and sentiment analysis with VADER to find aspects of a smartphone and its sentiment classification. From the 10 scenarios made in this reserach, it was found out three aspects that always appear, namely: the screen, camera, and battery aspects. It is concluded that these three aspects are the most important of a smartphone based on textual reviews.

Keywords— Topic Modeling; Sentiment Analysis; LDA; ecommerce

I. PENDAHULUAN

Smartphone di era modern ini telah menempuh perjalanan panjang untuk menjangkau sampai di tahun 2021, dan telah berubah secara drastis di sepanjang jalan. Ini adalah evolusi yang telah menggemparkan pasar. Dalam dekade terakhir, resolusi layar *smartphone* meningkat 11 kali lipat [1], pemimpin pasar tradisional seperti merek Nokia keluar dari pasar *smartphone* sementara merek-merek inovatif seperti Apple, Samsung, Huawei mendirikan ekosistem bisnis baru di pasar. Dengan cepatnya perkembangan teknologi baru, memahami tren pasar telah menjadi tugas yang semakin sulit. Dalam keadaan ini, ulasan produk online atau bisa disebut

sumber daya gratis yang mencerminkan sentimen konsumen tentang produk, saat ini menjadi perhatian.

Ulasan *online* adalah salah satu jenis sumber daya buatan pengguna yang populer dan penting, tempat pengguna dapat memublikasikan pengalaman dan pendapat mereka tentang produk, acara, atau layanan. Seiring aktivitas online yang terus berkembang, peran ulasan *online* menjadi semakin penting, terutama dalam proses pengambilan keputusan terkait tindakan pengguna di sistem online [2], [3].

Beberapa penelitian dalam [4], [5], dan [6] secara terus menerus menunjukkan bahwa kuantitas dan kualitas ulasan yang dibuat konsumen memiliki hasil yang berpengaruh positif pada niat pembeli dan menjadi sumber informasi utama bagi konsumen dan pemasar terkait kualitas produk. Namun, banyaknya ulasan *online* saat ini membuat proses penggalian informasi yang bermanfaat menjadi semakin sulit. Analisis ulasan online dapat membantu penjual dalam memahami minat dan keinginan konsumen sebelum meluncurkan produk baru.

Berdasarkan permasalahan di atas, penelitian ini mencoba untuk mencari tahu aspek yang berpengaruh dalam ulasan *online*. *Topic modeling* akan digunakan untuk menentukan aspek-aspek suatu merek *smartphone*. Selanjutnya akan dilakukan analisis sentimen terhadap aspek tersebut, sehingga dapat dikenali opini hasil ulasannya. *Dataset* untuk ulasan *online* ini akan diambil pada *website* Amazon dengan merek Apple, Samsung, Oneplus, dan Google. Dengan demikian, hal yang akan menjadi tujuan utama dalam penelitian ini adalah untuk mengimplementasikan penggabungan *topic modeling* guna menemukan aspek yang berpengaruh atau paling banyak dibicarakan pada produk *smartphone* dengan analisis sentimen dalam ulasan *online*

sehingga diperoleh perbandingan jumlah komentar bernuansa positif atau negatif untuk setiap aspek berpengaruh tersebut.

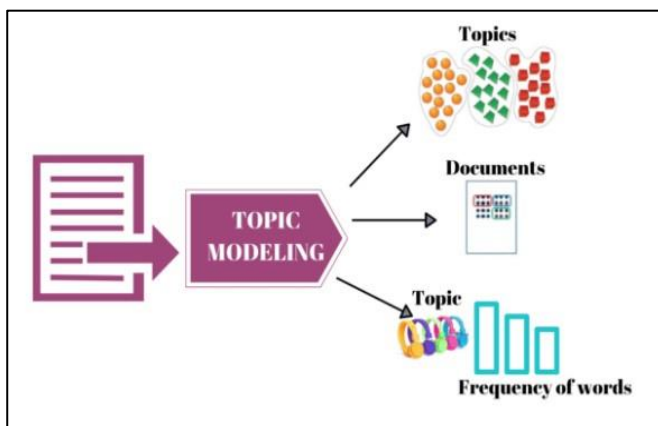
II. KAJIAN LITERATUR

A. Penelitian Terkait

Pada bagian ini akan dibahas beberapa teori yang berhubungan dengan penelitian yang dilakukan. Pada penelitian [7] sudah dilakukan analisis sentimen berbasis aspek pada *website book hotel* di *website Tripadvisor* dengan menggunakan LDA. Penelitian pada [8] menunjukkan analisis sentimen dengan menggunakan *Valence Aware Dictionary* dan *sEntiment Reasoner (VADER)*. VADER memiliki performa analisis sentimen yang sangat baik, dengan nilai *F1 Classification* = 0,96 dan *Accuracy* = 0,84, masing-masing) pada teks di sosial media. Penelitian [9] menerapkan topic modeling berdasarkan latent dirichlet allocation (LDA) untuk menganalisis artikel-artikel yang terdapat pada jurnal *Organization Research Method (ORM)* dan didapatkan sebanyak 15 topik dari 421 artikel yang cukup sesuai dari kajian yang dilakukan oleh manusia.

B. Topic Modeling

Dalam pembelajaran mesin dan pemrosesan bahasa alami, model topik adalah model generatif, yang menyediakan kerangka kerja probabilistik. Metode pemodelan topik umumnya digunakan untuk mengatur, memahami, mencari, dan meringkas arsip elektronik besar secara otomatis. "Topik" menandakan hubungan variabel yang tersembunyi, diperkirakan, yang menghubungkan kata-kata dalam kosakata dan kemunculannya dalam dokumen. Dokumen dipandang sebagai campuran topik. *Topic modeling* menemukan tema tersembunyi di seluruh koleksi dan membuat anotasi dokumen sesuai dengan tema tersebut. Setiap kata dipandang diambil dari salah satu topik tersebut. Sebuah distribusi cakupan dokumen topik dibuat dan memberikan cara baru untuk mengeksplorasi data pada perspektif topik [11]. Konsep *Topic modeling* ditunjukkan pada Gambar 1.



Gambar 1. *Topic Modeling*

C. Latent Dirichlet Allocation (LDA)

Dalam *NLP*, *Latent Dirichlet Allocation (LDA)* adalah teknik pemodelan topik yang secara otomatis menemukan topik dalam dokumen teks. LDA menganggap dokumen sebagai campuran dari berbagai topik dan setiap kata termasuk dalam salah satu topik dokumen. Algoritma ini pertama kali disajikan sebagai model grafis untuk penemuan topik [11]. LDA membayangkan serangkaian topik yang tetap. Setiap topik mewakili sekumpulan kata. Tujuan LDA adalah memetakan semua dokumen ke topik sedemikian rupa, sehingga kata-kata dalam setiap dokumen sebagian besar terkait dengan topik tersebut.

D. Analisis Sentimen dengan VADER

Analisis sentimen, juga disebut penambangan opini, adalah bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya. Ini mewakili ruang masalah yang besar. Ada juga banyak nama dan tugas yang sedikit berbeda, misalnya, analisis sentimen, penambangan opini, ekstraksi pendapat, penambangan sentimen, analisis subjektivitas, analisis pengaruh, analisis emosi, penambangan review, dll. Namun, semuanya sekarang berada di bawah bagian analisis sentimen atau penambangan opini [12].

VADER termasuk dalam jenis analisis sentimen yang didasarkan pada leksikon kata-kata yang berhubungan dengan sentimen. Dalam pendekatan ini, setiap kata dalam leksikon dinilai apakah itu positif atau negatif, dan dalam banyak kasus, seberapa positif atau negatif. Pada Tabel I adalah salah satu contoh kutipan dari leksikon Vader, di mana semakin positif kata tersebut maka peringkat positif lebih tinggi dan semakin negatif kata tersebut maka peringkat negatifnya semakin lebih rendah.

TABEL I
RATING SENTIMEN VADER

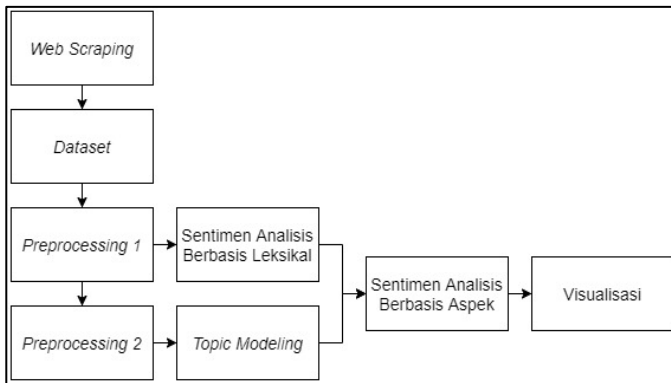
Word	Rating sentiment
<i>tragedy</i>	-3,4
<i>rejoiced</i>	2,0
<i>insane</i>	-1,7
<i>disaster</i>	-3,1
<i>Great</i>	3,1

VADER menganalisis sepotong teks yang diperiksa untuk melihat apakah ada kata-kata dalam teks yang ada dalam leksikon. Misalnya, kalimat "*the atmosphere was great and the people rejoiced*" memiliki dua kata dalam leksikon (baik dan bagus) dengan peringkat masing-masing 3,1 dan 2,0. VADER menghasilkan empat metrik sentimen dari suatu kalimat, yaitu positif, netral, negatif, dan skor *compound* yang mana adalah metrik terakhir hasil penjumlahan peringkat metrik sebelumnya yang akan di normalisasi ke

dalam nilai -1 (paling negatif), 0 (netral, dan lebih dari 0 sampai ke 1 (paling positif) [8].

III. METODOLOGI PENELITIAN & IMPLEMENTASI

Pada bagian ini akan dijelaskan tentang metodologi yang akan digunakan dalam penyusunan penelitian ini. Gambar 2 menunjukkan diagram bagaimana dilakukan metode untuk mencari aspek sentimen yang penting dari produk *smartphone*.



Gambar 2. Diagram Metodologi Penelitian

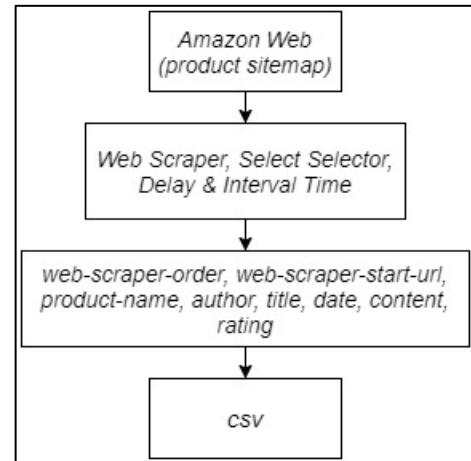
Proses pertama dalam penelitian ini adalah mengambil data, sumber data yang dipakai di penelitian ini berasal dari website *e-commerce* Amazon dengan tipe produk *smartphone*. Pengambilan data dilakukan dengan cara *scraping* melalui aplikasi *web scraping* pada *browser chrome*, setelah itu data disimpan dalam bentuk *csv*. Setelah dataset terkumpul, dilakukan *preprocessing* pada *dataset* yang akan menjadi input pada proses *topic modelling* dan analisis sentimen. *Topic modeling* dengan algoritma LDA digunakan untuk mencari aspek-aspek yang terkandung pada ulasan *smartphone*. Setelah mendapat aspek-aspek yang dominan dilakukan analisis sentimen dengan *VADER* sehingga mendapatkan analisis sentimen berbasis aspek-aspek yang ditemukan. Tahap terakhir adalah visualisasi hasil analisis sentimen berbasis aspek pada tiap merek *smartphone*.

A. Sumber Data

Data yang digunakan adalah produk *smartphone* dari merek Apple, Samsung, Google, dan Oneplus yang diambil dari *web e-commerce* Amazon menggunakan aplikasi *add-on Web scraper* dari *webscraper.io* yang di instalasi pada *browser Chrome*. Gambar 3 menunjukkan alur diagram *web scraping*, dimulai dari membuat *sitemap* untuk *web amazon*, membuat dan memilih *selector* untuk mengambil data yang diinginkan pada *start-url* yang dibuat, lalu menentukan *parameter* jeda waktu *parameter* dan *interval* saat ingin melakukan proses *scraping*. Setelah semua data selesai di *scraping*, Hasilnya akan disimpan dalam bentuk *csv*.

Jumlah data yang diambil berjumlah 51.782 *data entry* yang mencakup 4 merek, yaitu *Apple*, *Samsung*, *Google* dan

Oneplus dari model keluaran 2016 sampai 2019. Pada merek *Apple* ada 11 tipe yang diambil, yaitu *iPhone 7, 7 Plus, 8, 8 Plus, X, XR, XS, 11, 11 Pro, dan 11 Pro Max*. Merek *Samsung* diwakili dengan 14 tipe, yaitu *Galaxy Note 8, 9, 10, 10 lite, 10+, 20, Galaxy S7 Edge, S8, S8+, S9, S0, S20, dan S20 Ultra*. Untuk merek *Google* ada 7 tipe, yaitu *Pixel 2 XL, 3 3a, 3 XL, 4, 4a, dan Pixel 5*. Dan terakhir merek *OnePlus* yang diambil sebanyak 4 tipe, yaitu *OnePlus 7 Pro, 7T, 8, dan 8 Pro*.



Gambar 3. Diagram Web Scraping

Kode pada *Listing 1* memberikan penerapan *web scraping* dalam bentuk *json*. Nilai yang mengisi *_id* adalah nama dari *sitemap* yang dibuat, sedangkan *start Url* adalah *link* ulasan produk *smartphone* pada *website* Amazon yang ingin diambil. Dalam bagian *selector* ada *id: Product name* yang bertugas untuk mengambil nama produk *smartphone*. Nilai dalam *id: Review wrapper* adalah *selector element* untuk memilih daftar *item* apa saja yang ingin diambil.

Setelah itu ada *id: author* sebagai *child element* dari *id: Review wrapper* yang bertugas untuk mengambil nama pembeli *smartphone* tersebut, diikuti dengan *id: title* yang merupakan judul ulasan yang diberikan oleh *author id date* dimana lokasi dan tanggal ulasan tersebut terbit, *id: content* yang mana isi dari ulasan pengguna terhadap produk *smartphone* tersebut, dan *id: rating* adalah nilai bintang (1-5) yang diberikan oleh *author*. Pada *id: Click Next* adalah *selector* terakhir yang berfungsi untuk melakukan navigasi pada halaman ulasan berikutnya, lalu kembali lagi dengan proses yang sama dimulai dari *id: Product name* sampai ulasan yang ada diambil.

```
{
  "_id": "amazon-reviews-scraper-2020",
  "startUrl": ["https://www.amazon.com/Ovente-Dual-Sided-Magnification-Electrical-MPWD3185BZ1X7X/product-reviews/B074GCRS9D"],
  "selectors": [{"id": "Product name",
    "type": "SelectorText",
    "parentSelectors": ["_root"],
    "selector": "div[class*='product title']"},
    {"multiple": false,
```

```

    "regex": "",
    "delay": 0
  }, { "id": "Review wrappers",
    "type": "SelectorElement",
    "parentSelectors": ["_root", "Click Next"],
    "selector": "div.a-section.review",
    "multiple": true,
    "delay": 0
  }, { "id": "author",
    "type": "SelectorText",
    "parentSelectors": ["Review wrappers"],
    "selector": "span.a-profile-name",
    "multiple": false,
    "regex": "",
    "delay": 0}, { "id": "title",
    "type": "SelectorText",
    "parentSelectors": ["Review wrappers"],
    "selector": "a.a-size-base.review-title",
    "multiple": false,
    "regex": "",
    "delay": 0
  },
  { "id": "date",
    "type": "SelectorText",
    "parentSelectors": ["Review wrappers"],
    "selector": "span.a-size-base.a-color-
secondary",
    "multiple": false,
    "regex": "",
    "delay": 0
  }, { "id": "content",
    "type": "SelectorText",
    "parentSelectors": ["Review wrappers"],
    "selector": "div.a-row.review-data span.a
size-base",
    "multiple": false,
    "regex": "",
    "delay": 0
  }, { "id": "rating",
    "type": "SelectorText",
    "parentSelectors": ["Review wrappers"],
    "selector": "span.a-icon-alt",
    "multiple": false,
    "regex": "",
    "delay": 0
  }, { "id": "Click Next",
    "type": "SelectorElementClick",
    "parentSelectors": ["_root"],
    "selector": "div.review-views",
    "multiple": false,
    "delay": "4500",
    "clickType": "clickMore",
    "discardInitialElements": "discard",
    "clickElementUniquenessType": "uniqueText"
  ]
}

```

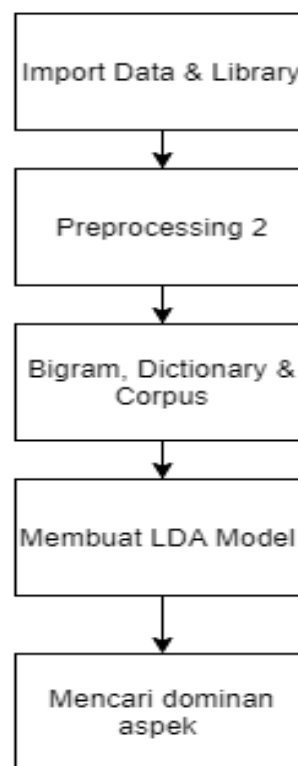
Listing 1. Web Scraping

B. Topic Modeling

Untuk mencari aspek-aspek pada produk *smartphone*, digunakan *topic modeling*, yang menggunakan algoritma *LDA* menggunakan *library*. Gambar 4 menjelaskan diagram alur untuk mendapatkan aspek suatu *smartphone*. Pada percobaan ini digunakan 10 jenis kondisi data sebagai input sebagaimana yang tertera pada Tabel II.

TABEL II.
JENIS SKENARIO DATASET

Merek	Kondisi
Semua Merek	Noun Tag, Noun & Adj Tag
Apple	Noun Tag, Noun & Adj Tag
Samsung	Noun Tag, Noun & Adj Tag
OnePlus	Noun Tag, Noun & Adj Tag
Google	Noun Tag, Noun & Adj Tag



Gambar 4. Diagram Topic Modeling

C. Preprocessing

Setelah didapatkan 10 jenis kondisi data sebagai input, maka dilakukan *pre-processing* yang kedua dengan membuat semua teks menjadi huruf kecil (*lowercase*), menghapus teks dalam kurung siku, menghapus tanda baca, dan menghapus kata-kata yang mengandung angka. Setelah *data cleaning* dilakukan, setiap ulasan pada *dataframe* akan diubah menjadi bentuk *list* dan dilakukan tokenisasi untuk merubah bentuk kalimat menjadi menjadi kumpulan kata-kata berbentuk *list*.

TABEL III

HASIL PROSES *PREPROCESSING STOPWORDS* DAN *LEMMATIZATION (NOUN)*

Sebelum Proses	Sesudah Proses
'an', 'amazing', 'purchase', 'gamechanger', ['the', 'phone', 'came', 'as', 'described', 'in', 'cosmetically', 'it', 'had', 'the', 'signs', 'of', 'good', 'use', 'with', 'couple', 'of', 'scuffs',	'purchase', 'gamechanger', ['packaging', 'cardboard', 'box', 'work', 'flawless', 'couple', 'scuff'

Proses selanjutnya dilakukan proses penghilangan stopwords, membuat *bigram*, dan melakukan *lemmatization* dengan kondisi hanya kata benda (*NOUN*) dan kata benda (*NOUN*) dengan sifat (*ADJ*) saja. Hasil proses ini bisa dilihat pada Tabel III yang menunjukkan hanya kata benda saja yang diambil.

D. Membuat input LDA (Dictionary & Corpus)

Untuk membuat model LDA dibutuhkan 2 input yaitu, *dictionary (id2word)* dan korpus dari proses sebelumnya. Dalam membuat dua bentuk *input* ini dilakukan dengan bantuan *library Gensim* dimana *library* ini membuat id unik untuk setiap kata pada dokumen beserta frekuensi kata tersebut di tiap dokumen. Contoh seperti pada Gambar 5, dimana (0,2) menjelaskan 0 adalah id *word* yang muncul dua kali pada dokumen tersebut. Pada *output* selanjutnya bisa dilihat 0 adalah id untuk kata 'condition' di dokumen tersebut. Bentuk seperti itu yang akan dipakai sebagai input membuat model LDA. Pada bagian listing 2 adalah bagian untuk membuat *input model LDA* dengan bantuan *Gensim*.

```
# Create Dictionary
id2word = corpora.Dictionary(data_lemmatized)
# Create Corpus
texts = data_lemmatized
# Term Document Frequency
corpus = [id2word.doc2bow(text) for text in
texts]
# View
print(corpus[:1])
# Human readable format of corpus (term-
frequency)
[[id2word[id], freq] for id, freq in cp] for cp
in corpus[:1]]
```

Listing 2. Membuat Input untuk Model LDA

```
[[[0, 2), (1, 1), (2, 1)]]
[ ] # Human readable format of corpus (term-frequency)
[[[id2word[id], freq] for id, freq in cp] for cp in corpus[:1]]
[[['condition', 2), ('ding', 1), ('side', 1)]]
```

Gambar 5. Hasil Pembentukan input LDA

E. Membuat LDA Model

Pada proses ini dilakukan pembuatan model LDA yang menggunakan input *dictionary* dan korpus yang sudah dibuat tadi. Parameter yang ditentukan dalam membuat model LDA terdapat pada Gambar 6. Parameter *alpha* adalah sebuah *hyperparameter* yang mempengaruhi ketersebaran topik, *chunksize* adalah seberapa banyak dokumen yang akan dipakai pada setiap *training*, *update_every* yang akan mempengaruhi seberapa sering parameter model akan diperbarui dan *passes* adalah parameter untuk menentukan seberapa banyak *training* yang dilakukan. Setelah model jadi akan dihitung nilai koherensinya untuk melihat seberapa baik model yang sudah dibuat.

Setelah proses model dibuat akan dicari jumlah topik yang paling optimal dengan cara membuat beberapa model LDA dengan jumlah topik yang berbeda dan memilih jumlah topik yang memiliki nilai koherensi tinggi. Koherensi topik adalah ukuran yang digunakan untuk mengevaluasi model topik: metode yang secara otomatis menghasilkan topik dari kumpulan dokumen, menggunakan model variabel laten. Setiap topik yang dihasilkan terdiri dari kata-kata, dan koherensi topik diterapkan ke N kata teratas dari topik tersebut. Ini didefinisikan sebagai rata-rata / median dari skor kemiripan kata berpasangan dari kata-kata dalam topik tersebut. Model yang baik akan menghasilkan topik yang koheren, yaitu topik dengan skor koherensi topik yang tinggi. Topik yang baik adalah topik yang dapat dideskripsikan dengan label pendek [13].

```
Building the Topic Model
# Build LDA model
lda_model = gensim.models.ldamodel.LdaModel(corpus=corpus,
id2word=id2word,
num_topics=20,
random_state=100,
update_every=1,
chunksize=100,
passes=10,
alpha='auto',
per_word_topics=True)
```

Gambar 6. Parameter Model LDA

F. Analisis Sentimen Berbasis Aspek

Setelah aspek-aspek suatu *smartphone* didapatkan dengan bantuan *LDA topic modeling*, maka akan dilakukan proses analisis sentimen pada kolom 'review' yang ada pada *dataframe* hasil *topic modeling* dengan bantuan *library VADER*. Pada tahap ini tidak dilakukan *pre-processing* lebih lanjut dikarenakan *VADER* dapat memberikan skor sentimen yang berbeda pada jenis teks yang ada. Tabel IV menunjukkan hasil akhir analisis sentimen terhadap aspek yang dominan untuk tiap ulasan yang diberikan oleh *author*, sehingga bisa dilihat sentimen untuk tiap ulasan berdasarkan aspeknya. *Dataframe* yang sudah jadi, disimpan dalam bentuk csv untuk selanjutnya divisualisasikan.

TABEL IV
HASIL ANALISIS SENTIMEN BERBASIS ASPEK

Model	Year	Review comments	Rate	Aspect	Sent
iPhone 7 Plus	2019	phone came in good condition came with everyth...	5	screen	+
iPhone 7 Plus	2020	excelente producto..	1	work	-
iPhone 7 Plus	2018	This phone came as good and screen...	5	screen	+
iPhone 7 Plus	2018	little to no scratches looks almost brand new ...	4	battery	+
iPhone 7 Plus	2018	got it yesterday looks almost new works well t...	5	speaker	+
iPhone 8 Plus	2019	So far, the phone is great...	4	condition	+
iPhone 8 Plus	2020	LOVE IT. Looks amazing. Refurbishe d great. One little scuff on the top right but that's ok with me...	4	battery	+
iPhone 8 Plus	2020	Turned right on. Delivered quickly and packaged perfect. Came with wall peice and...	5	work	+
iPhone 8 Plus	2019	Overall, great phone fast Delivery and ...	5	condition	-

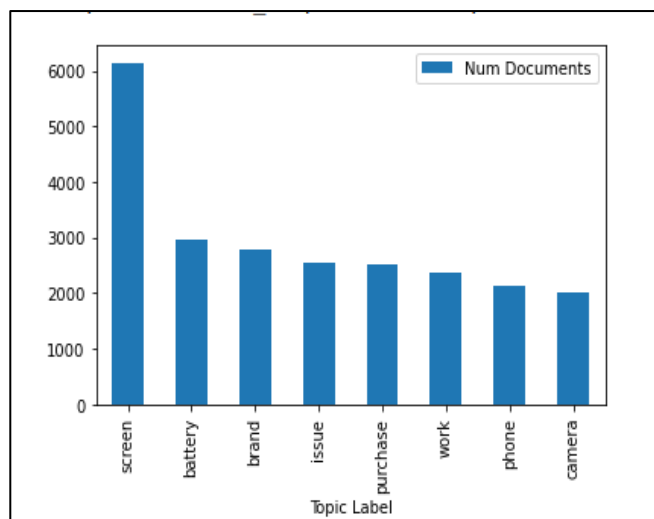
IV. HASIL EKSPERIMEN DAN EVALUASI

A. Evaluasi Aspek dengan Noun Tag

Pada bagian ini akan dibahas hasil eksperimen dalam bentuk visualisasi dari dataset *smartphone* untuk semua merek maupun untuk masing-masing merek *smartphone* dengan *Part of Speech (POS) Tag* kata benda.

A.1 Evaluasi Aspek Apple Smartphone Noun Tag

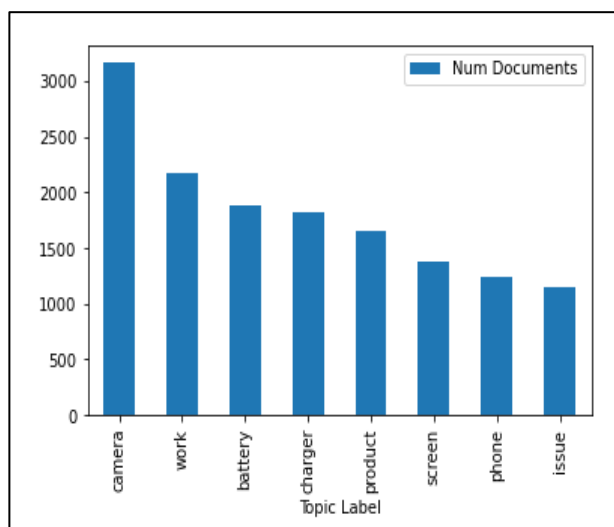
Gambar 7 menunjukkan bahwa aspek *screen* adalah aspek yang paling dominan pada dokumen ulasan *smartphone* Apple diikuti dengan aspek baterai, *brand*, *issue*, *purchase*, *work*, *phone*, dan *camera*.



Gambar 7. Aspek Tertinggi pada Apple

A.2 Evaluasi Aspek Samsung Smartphone Noun Tag

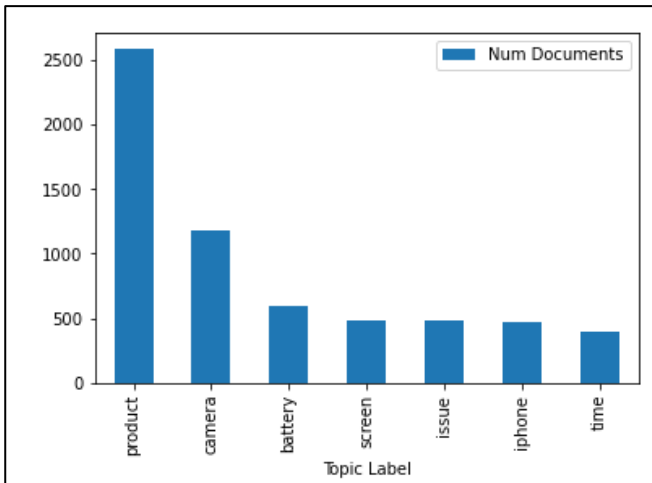
Gambar 8 menunjukkan bahwa *camera* adalah aspek yang paling dominan pada dokumen ulasan *smartphone* Samsung diikuti dengan aspek *work*, *battery*, *charger*, *product*, *screen*, *phone*, dan *issue*.



Gambar 8. Aspek Tertinggi pada Samsung

A.3 Evaluasi Aspek OnePlus Smartphone Noun Tag

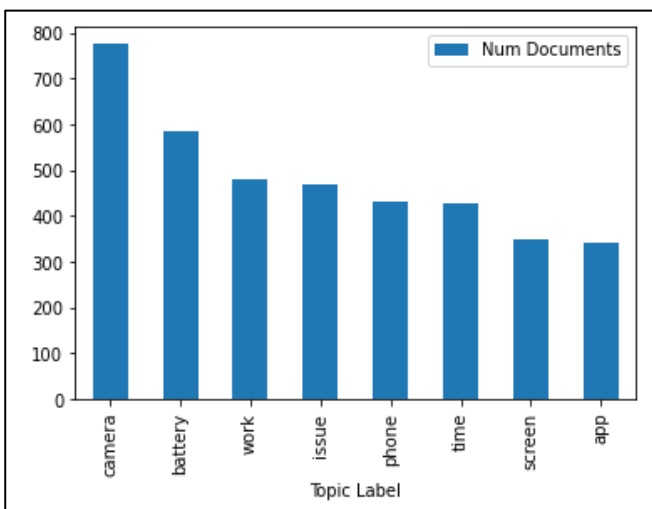
Gambar 9 menunjukkan bahwa product adalah aspek yang paling dominan pada dokumen ulasan smartphone OnePlus diikuti dengan aspek camera, battery, screen, issue, iphone, dan time.



Gambar 9. Aspek Tertinggi pada OnePlus

A.4 Evaluasi Aspek Google Smartphone Noun Tag

Gambar 10 menunjukkan bahwa camera adalah aspek yang paling dominan pada dokumen ulasan smartphone Google diikuti dengan aspek battery, work, issue, phone, time, screen, dan app.



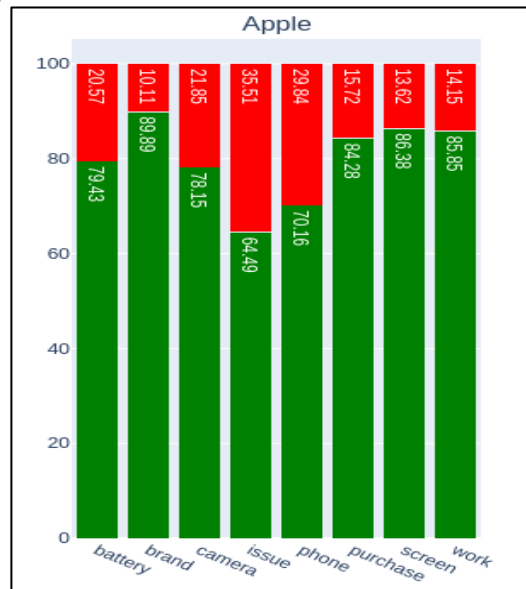
Gambar 10. Aspek Tertinggi pada Smartphone Google

B. Analisis Sentimen Berbasis Aspek Pada Merek Apple

Pada Gambar 11 sampai 14, grafik berwarna hijau menunjukkan sentimen positif sedangkan grafik berwarna merah menunjukkan sentimen negatif. Informasi pada sumbu

x menunjukkan merek smartphone dan sumbu y menunjukkan jumlah presentase analisis sentimen secara keseluruhan.

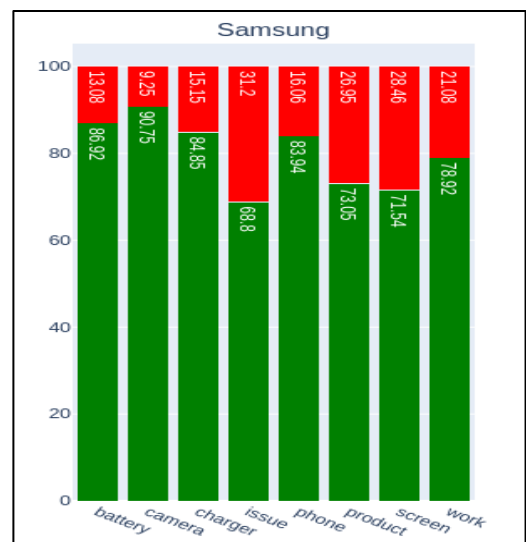
Pada Gambar 11 ditunjukkan hasil visualisasi analisis sentimen terhadap aspek merek Apple. Bisa dilihat aspek brand yang memiliki sentimen positif paling tinggi, sedangkan aspek issue yang memiliki sentimen negatif paling tinggi.



Gambar 11. Analisis Sentimen Berbasis Aspek pada Apple

B.1 Analisis Sentimen Berbasis Aspek Pada Merek Samsung

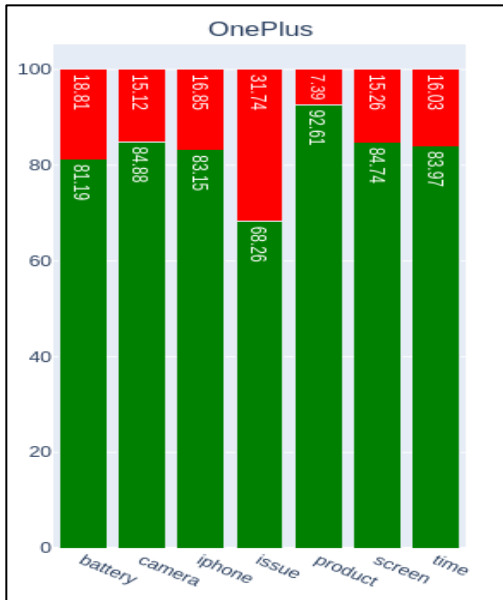
Pada Gambar 12 ditunjukkan hasil visualisasi analisis sentimen terhadap aspek merek Samsung. Bisa dilihat aspek kamera yang memiliki sentimen positif paling tinggi, sedangkan aspek issue dan layar yang memiliki sentimen negatif paling tinggi.



Gambar 12. Analisis Sentimen Berbasis Aspek pada Samsung

B.2 Analisis Sentimen Berbasis Aspek Pada Merek OnePlus

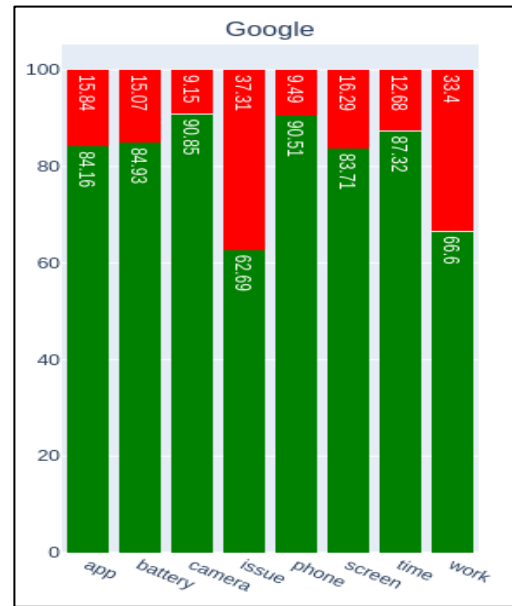
Pada Gambar 13 ditunjukkan hasil visualisasi analisis sentimen terhadap aspek merek OnePlus. Bisa dilihat aspek *product* yang memiliki sentimen positif paling tinggi, sedangkan aspek *issue* memiliki sentimen negatif paling tinggi.



Gambar 13. Analisis Sentimen Berbasis Aspek pada OnePlus

B.3 Analisis Sentimen Berbasis Aspek Pada Merek Google

Pada Gambar 14 ditunjukkan hasil visualisasi analisis sentimen terhadap aspek merek Google. Bisa dilihat aspek kamera dan *phone* yang memiliki sentimen positif paling tinggi, sedangkan aspek *issue* memiliki sentimen negatif paling tinggi.



Gambar 14. Analisis Sentimen Berbasis Aspek pada Google

C. Evaluasi Hasil

Tabel V menunjukkan aspek-aspek yang didapatkan dari 10 skenario penelitian ini, sedangkan Tabel VI adalah potongan contoh dari tabel frekuensi kemunculan aspek terhadap 10 skenario pengambilan data yang ada. Dari 10 skenario yang dilakukan didapatkan 25 aspek unik yang muncul. Tiga aspek teratas yang selalu muncul dari skenario penelitian yang telah dilakukan adalah *screen*, *battery* dan *camera*.

TABEL V
ASPEK UNIK YANG DIDAPKANT DARI 10 SKENARIO

Aspek Unik
<i>screen, battery, brand, issue, purchase, work, phone, camera, charger, product, iphone, time, app, day, great, problem, unlocked, review, samsung, good, pro,feature, android, condition, love</i>

TABEL VI
FREKUENSI ASPEK YANG MUNCUL

Skenario	Aspek				
	Screen	Battery	Brand	Issue	Camera
Apple noun	1	1	1	1	1
Samsung noun	1	1	0	1	1
OnePlus Noun	1	1	0	1	1
Google Noun	1	1	0	1	1
Apple noun adj	1	1	1	0	1
Samsung noun adj	1	1	0	0	1
Oneplus noun adj	1	1	0	0	0
Google noun adj	1	1	0	0	0
All brand noun	1	1	0	1	1
All brand adj	1	0	0	1	1
Total aspek	10	9	2	6	8

Bisa dilihat aspek *screen* selalu ada di tiap skenario yang dilakukan diikuti oleh aspek *battery* dan *camera*. Dengan kata lain ketiga aspek tersebut dapat dianggap sebagai aspek yang paling berpengaruh pada *smartphone* dalam *dataset* penelitian ini. Jika digabungkan dengan analisis sentimen, pengguna maupun penggiat lainnya dapat memperoleh *insight* yang bermanfaat yang bisa dipakai untuk mengambil keputusan atau penelitian selanjutnya terkait dengan pemanfaatan ataupun pemasaran *smartphone*.

V. KESIMPULAN DAN PENGEMBANGAN

Implementasi *topic modeling* menggunakan algoritma LDA sudah berhasil dibuat untuk mencari aspek suatu *smartphone*. Pada penelitian ini didapatkan 3 aspek teratas yang selalu muncul dari skenario yang dilakukan yaitu *screen*, *battery*, dan *camera*, sehingga bisa disimpulkan pada penelitian ini bahwa 3 aspek tersebut yang menjadi aspek paling berpengaruh dalam suatu *smartphone*. Implementasi analisis sentimen juga sudah berhasil dilakukan sehingga menghasilkan analisis sentimen terhadap aspek-aspek suatu *smartphone*. Ditemukan dalam skenario *tagging* kata benda

dan kata sifat, pengguna sering menggunakan kata kunci atau aspek '*great*' dalam memberikan sentimen positif mereka terhadap *smartphone* yang diulas.

Untuk penelitian selanjutnya diusulkan untuk membangun model yang lebih spesifik atau menggunakan model yang dikembangkan secara langsung dengan penggabungan aspek dengan sentimennya (*multi-classification*). Pada proses *automatic topic labeling* juga bisa ditingkatkan dengan cara yang lain seperti menggunakan *graph-based* selain dengan menggunakan *keyword* tertinggi dari output *LDA topic modeling*. Diharapkan bahwa metoda kerja yang diusulkan melalui riset ini dapat bermanfaat untuk keperluan pemasaran produk selain *smartphones*.

DAFTAR PUSTAKA

- [1] Bryant, Laurence. (2019) Ten Years of Mobile Evolution: A Look at the Samsung Galaxy S Series. [Online]. Tersedia: <https://community.arm.com/developer/ipproducts/processors/b/processors-ip-blog/posts/samsung-galaxy-sten-years-of-mobile-evolution>.
- [2] Nielsen. Online. (2015) 81 percent of online holiday shoppers read online customer reviews, according to Nielsen online. [Online]. Tersedia: <http://www.nielsen-online.com/pr/pr-081218.pdf>.
- [3] A. Palmer. (2015) Web shoppers trust customer reviews more than friends. [Online]. Tersedia: <http://www.adweek.com/news/advertising-branding/web-shoppers-trust-customer-reviews-more-friends-100313>.
- [4] D. H. Park, J. Lee, & I. Han, "The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement," *International Journal Electronic Commerce*, vol. 11, no. 4, pp.125–148, July 2007.
- [5] Hu, N., Liu, L. & Zhang, J.J. "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects," *Information Technology Management*, vol. 9, 201–214, 2008.
- [6] Park, D.H, Kim, S, "The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews," *Electronic Commerce Research and Application*, vol. 7, no. 4, pp. 399–410, December. 2008
- [7] Akhtar, N., Zubair, N., Kumar, A. & Ahmad, T, "Aspect based Sentiment Oriented Summarization of Hotel Reviews," *Procedia Computer Science*, vol. 115, pp. 563–571. Elsevier, 2017.
- [8] Hutto, C. & E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," *Prosiding International AAAI Conference on Web and Social Media*, 2014, vol. 8, no. 1.
- [9] A. Gaur, "Topic Models As A Novel Approach To Identify Themes In Content Analysis," *Prosiding Academy of Management Proceedings*, AOM 2017.
- [10] M. Blei., A. Ng., & M. Jordan, "Probabilistic Topic Model," *Communications of the ACM* vol.55, no. 4, 2012.
- [11] M. Blei, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp.993-1022., 2003.
- [12] B. Liu, *Sentiment Analysis and Subjectivity Handbook of Natural Language Processing*, Cambridge, United Kingdom: Cambridge University Press, 2015.
- [13] D. Newman, J.H. Lau, K. Grieser, & T. Baldwin, "Automatic evaluation of topic coherence," *Prosiding Human Language Technologies- North American Chapter of the Association for Computational Linguistics*, 2010, pp. 100-108.