

Analisis Sentimen pada Ulasan Handphone dengan Algoritma FP-Growth

Roza Marmay^{*1}, Leony Lidya², Sarjon Defit³

¹Politeknik ATI Padang, Indonesia

^{2,3}Universitas UPI "YPTK" Padang, Indonesia

Email: ¹rozamarmay@politeknikatipdg.ac.id, ²leonylidya@yahoo.com, ³sarjond@yahoo.co.uk

Abstrak

Masyarakat dewasa ini lebih sering melakukan pembelian produk melalui daring. Hal ini tentu membutuhkan kecermatan dalam membaca ulasan guna mendapatkan produk yang baik. Sentimen analisis dilakukan untuk menganalisa ulasan pengunjung dari komentar sebuah produk dalam media sosial. Penelitian kali ini difokuskan kepada ulasan produk handphone merk asus zenfone2 yang diambil dari amazon.com guna mengetahui sentimen pengunjung website terkait produk yang dipilih. Proses pengolahan data tersebut dimulai dari pemilahan ulasan yang didapat menjadi perkalimat untuk mempermudah proses selanjutnya. Untuk mendapatkan *noun* dan *adjective* dari kalimat tersebut, dilakukan tahapan *preprocessing* seperti *lower case*, tokenisasi, *lemmatization*, serta *POS tagging*. *Noun* yang didapat dari *preprocessing* tersebut digunakan dalam algoritma FP-growth untuk menemukan fitur pada handphone asus zenfone2 yang sering dibicarakan. Sedangkan *adjective* yang di dapat, akan digunakan untuk mendeteksi apakah kalimat yang mengandung fitur tersebut bernilai positif atau negatif. Hasil dari analisis ini dapat digunakan *customer* dalam mempertimbangkan produk tersebut tanpa harus membaca ulasan satu persatu.

Kata Kunci: Analisis Sentimen, FP-Growth, Opinion Words

Abstract

Nowadays, people more frequently purchase products online. This certainly requires careful attention in reading reviews to find quality products. Sentiment analysis is conducted to analyze customer reviews of a product on social media. This study focuses on reviews of the Asus Zenfone 2 mobile phone taken from Amazon.com to understand the sentiment of website visitors regarding the selected product. The data processing begins by segmenting the reviews into individual sentences to facilitate the next steps. To extract nouns and adjectives from these sentences, preprocessing steps such as lowercasing, tokenization, lemmatization, and POS tagging are performed. The nouns obtained from this preprocessing are used in the FP-growth algorithm to identify features of the Asus Zenfone 2 mobile phone that are frequently discussed. Meanwhile, the adjectives are used to detect whether the sentences containing these features are positive or negative. The results of this analysis can help customers consider the product without having to read each review individually.

Keywords: FP-Growth, Opinion Words, Sentiment Analysis

1. PENDAHULUAN

Perkembangan penggunaan internet sekarang ini semakin pesat. Semakin mudahnya penggunaan dan pengaksesan internet oleh semua pihak, membuat internet tidak hanya sebagai media mencari informasi tapi juga bertukar informasi. Sekarang ini juga banyak pengusaha memanfaatkannya sebagai media promosi produk serta sebagai sarana berbisnis seperti *online shopping*. Menurut data Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) Indonesia merupakan negara yang memiliki total populasi sebanyak 264 juta jiwa penduduk, ada sebanyak 171,17 juta jiwa atau sekitar 64,8 persen yang sudah terhubung ke internet (litbang, 2019).

Ulasan yang dituliskan oleh seseorang terhadap suatu produk dapat digunakan oleh pengguna internet lainnya sebagai acuan bahan pertimbangan untuk memilih suatu produk, sedangkan bagi perusahaan dapat dijadikan sebagai bahan evaluasi dan mengetahui selera pasar. Misalnya, sebelum membeli suatu produk, kebanyakan orang cenderung mencari melalui ulasan tentang produk tersebut

yang akan membantu mereka dalam membuat pilihan. Berdasarkan survey yang dilakukan oleh brightlocal, 79% konsumen mengatakan mereka mempercayai ulasan online lebih dari rekomendasi pribadi dari teman atau keluarga (brightlocal, 2020).

Analisis sentimen atau penggalian opini adalah studi yang menganalisis pendapat, pemikiran, dan kesan orang mengenai berbagai topik, subjek, dan produk atau layanan. Yang bertujuan untuk mengotomatisasi tugas mengidentifikasi opini dan sentimen yang orang-orang ungkapkan, dan kemudian mengklasifikasikan polaritas sentimennya (Vanaja, et al, 2018). Sejak awal tahun 2000, analisis sentimen telah berkembang menjadi salah satu area penelitian paling aktif dalam Natural Language Processing (Liu, 2020). Pengimplementasian analisis sentimen telah menyebar ke banyak domain, mulai dari bisnis, layanan kesehatan, pemilihan umum, analisis produk, hingga riset pasar.

Beberapa penelitian serupa telah dilakukan sebelumnya, diantaranya adalah melakukan analisa sentimen menggunakan naive bayes untuk mengkategorikan ulasan positif dan negatif dari pelanggan amazon atas kategori produk yang berbeda seperti ponsel, aksesoris, alat musik dan barang elektronik (hague, 2018). Penelitian tersebut mampu menghasilkan akurasi rata-rata di atas 90% dengan mencoba beberapa simulasi yang berbeda seperti cross validation, rasio data latih-uji, dan teknik ekstraksi fitur yang berbeda. Penelitian lainnya juga telah dilakukan terkait sentimen analisis dengan yang dilakukan terkait sentimen analisis berbasis tweet berbahasa indonesia menggunakan metode deep belief network (ira zulfa, et al, 2017). Penelitian terkait Ulasan Produk juga pernah dilakukan sebelumnya namun menggunakan Metode Naive Bayes (Billy Gunawan, et al, 2018), serta Ekstraksi Opini Berbasis Fitur Untuk Review Produk Berbahasa Indonesia Dengan Algoritma Fp-growth (Dhahyani Tetya Arum, et al, 2015).

Terdapat beberapa pendekatan yang umum digunakan untuk melakukan analisis sentimen, antara lain dengan menggunakan machine learning, berbasis leksikon, hybrid dan teknik lainnya seperti aspect-based approach, transfer learning, dan multimodal sentiment analysis (Birjali et al, 2021). Pada penelitian ini akan dilakukan analisis sentimen pada ulasan handphone dengan algoritma fp-growth. Adapun tujuan yang ingin dicapai adalah mengetahui fitur yang sering dibicarakan terkait asus zenfone 2 dan mengetahui ulasan tersebut bersifat positif atau negatif sehingga dapat dijadikan sebagai bahan pertimbangan sebelum melakukan pembelian produk tersebut.

2. METODE PENELITIAN

Dalam penelitian ini ada beberapa istilah yang digunakan sebagai pendukung proses Penelitian :

2.1. Data Mining

Data Mining menurut Turban, dkk merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakrit dari berbagai database besar/*Data Warehouse* (Tampubolon Kennedy, 2013).

2.2. Text Mining

Text mining merupakan salah satu cara dalam ilmu komputer untuk menentukan koneksi yang kuat dari teks yang ada dengan *natural language processing*, *Data Mining* atau machine learning, pengambilan keputusan dan manajemen pengetahuan (Radovanovic and Ivanovic, 2008). *Text mining* juga didefinisikan sebagai variasi dari *Data Mining* yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar (Kurniawan Bambang, 2012).

2.3. Text Preprocessing

Preprocessing adalah tahapan mengubah suatu dokumen kedalam format yang sesuai agar dapat diproses. Terdapat 3 tahapan preprocessing dalam penelitian ini, yaitu (Milatina et.al, 2012):

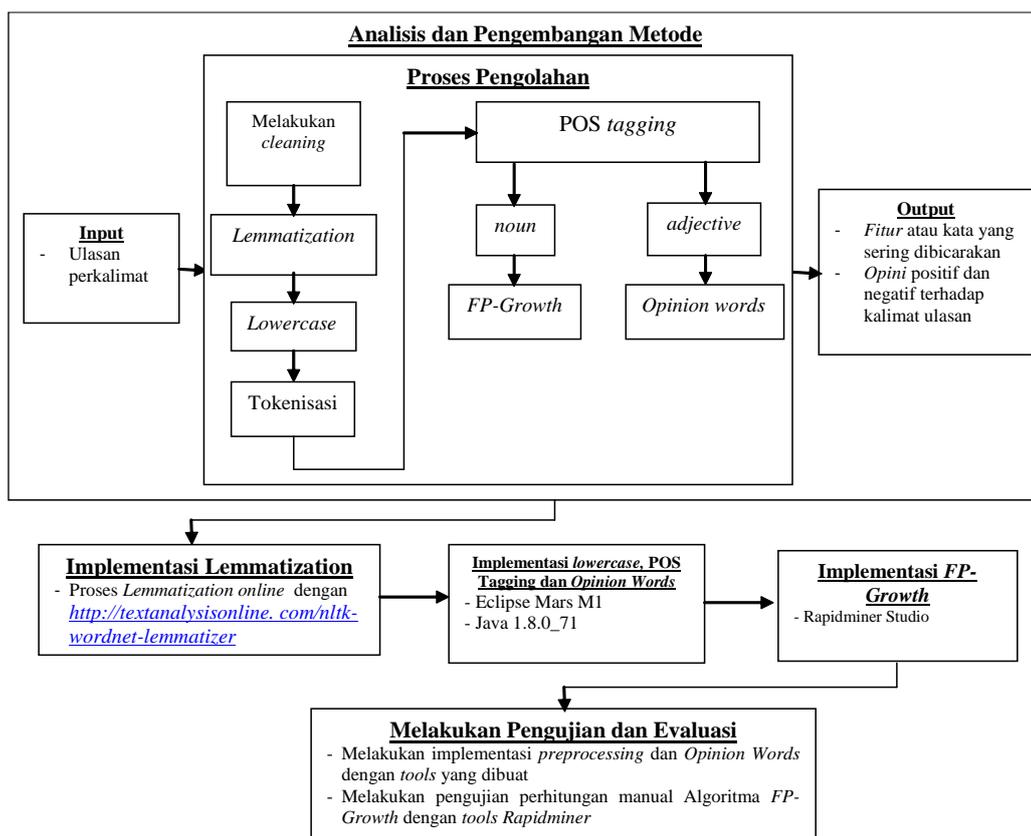
a. Tokenization, merupakan tahapan penguraian string teks menjadi term atau kata.

- b. Stopword Removal, merupakan tahapan penghapusan kata-kata yang tidak relevan dalam penentuan topik sebuah dokumen dan yang sering muncul pada sebuah dokumen, misal "and", "or", "the", "a", "an" pada dokumen berbahasa Inggris.
- c. Stemming, merupakan tahapan pengubahan suatu kata menjadi akar kata nya dengan menghilangkan imbuhan awal atau akhir pada kata tersebut. Selain dengan melakukan stemming, dapat juga dengan *lemmatization*.

2.4. Algoritma FP-Growth

Algoritma FP-Growth merupakan pengembangan dari algoritma apriori. Sehingga kekurangan dari algoritma apriori diperbaiki oleh algoritma FP-Growth. Pada algoritma *Apriori* memerlukan *generate candidate* untuk mendapatkan *frequent itemsets*. Akan tetapi, di algoritma *FP-Growth* *generate candidate* tidak dilakukan karena *FP-Growth* menggunakan konsep pembangunan *tree* dalam pencarian *frequent itemsets*. Hal tersebutlah yang menyebabkan algoritma *FP-Growth* lebih cepat dari pada algoritma *Apriori*. Karakteristik algoritma *FP-Growth* adalah struktur data yang digunakan berupa *tree* yang disebut dengan *FP-Tree* (Erwin, 2009)

Struktur data yang digunakan untuk mencari *frequent itemset* dengan algoritma *FP-Growth* adalah perluasan dari sebuah pohon prefix, yang biasa disebut FP-Tree. Dengan menggunakan FP-Tree, algoritma *FP-Growth* dapat langsung mengekstrak *frequent itemset* dari FP-Tree yang telah dibentuk dengan menggunakan prinsip *devided and conquer* (Ririanti, 2014).



Gambar 1 Tahapan Penelitian

Gambar 1 menampilkan tahapan penelitian secara umum yaitu proses *cleaning*, *lemmatization*, *lowercase*, tokenisasi dan postagging. Selanjutnya dilakukan *Lemmatization*, POS Tagging, dan FP-Growth.

Proses *cleaning* merupakan proses untuk membuang duplikasi data dan memperbaiki kesalahan pada data. *Lemmatization* Merupakan proses untuk mengembalikan kata menjadi bentuk kamusnya,

implementasi tahapan *lemmatization* dilakukan secara *online* dengan <http://textanalysisonline.com/nltk-wordnet-lemmatizer>. *Lowercase* Merupakan proses untuk mengecilkan semua huruf pada kalimat setelah proses *lemmatization*. Tokenisasi Merupakan proses untuk memilah kalimat menjadi perkata, untuk digunakan pada proses selanjutnya. *POS Tagging* untuk mengidentifikasi suatu kata, apakah kata tersebut merupakan *noun*, *adjective*, *verb*, *adverb*, dan lain sebagainya. Analisis yang dilakukan menggunakan proses *Data Mining* dengan Menggunakan Algoritma FP-Growth menggunakan Rapidminer, untuk mencari *frequent itemset* pada setiap kalimat ulasan. Kata yang diolah merupakan *noun* dari hasil yang telah didapat pada proses *pos tagging*. *Opinion words* yaitu untuk menentukan apakah ulasan tersebut bersifat positif atau negatif yang dilihat dari ekspresi kalimat tersebut.

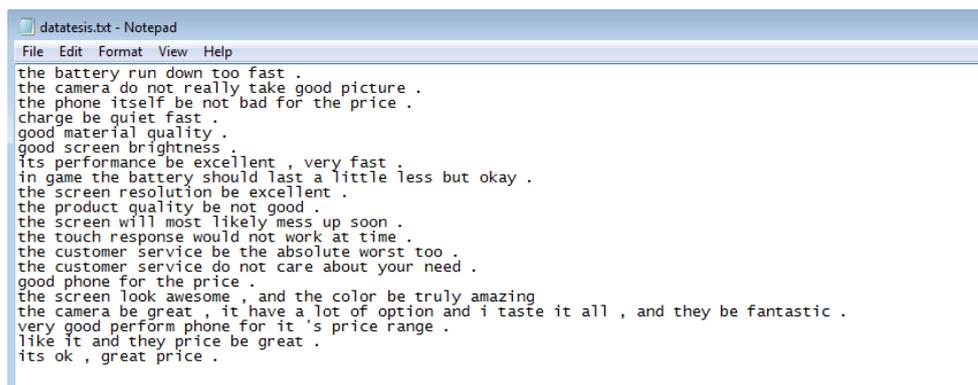
Pada proses ini yang dilakukan adalah implementasi *preprocessing* dengan *tools* yang penulis buat serta melakukan pengujian dari hasil pencarian secara manual dengan *Rapidminer*. Hasil pengujian digunakan untuk penarikan kesimpulan dari penelitian yang dilakukan.

Keluaran yang diharapkan dari proses yang dilakukan adalah *figure* handphone yang sering menjadi pembicaraan, dan positif dan negatif dari *opinion words* terhadap kalimat ulasan tersebut.

3. HASIL DAN PEMBAHASAN

3.1. Data Penelitian

Data yang digunakan bersumber dari data ulasan handphone yang di ambil dari www.amazon.com dengan jumlah data analisis sebanyak 20 transaksi dan 28 atribut, serta data uji sebanyak 420 transaksi dan 366 atribut. Gambaran umum data yang digunakan dapat dilihat pada gambar 2.



Gambar 2. Data Analisis

3.2. Proses Online Lemmatization

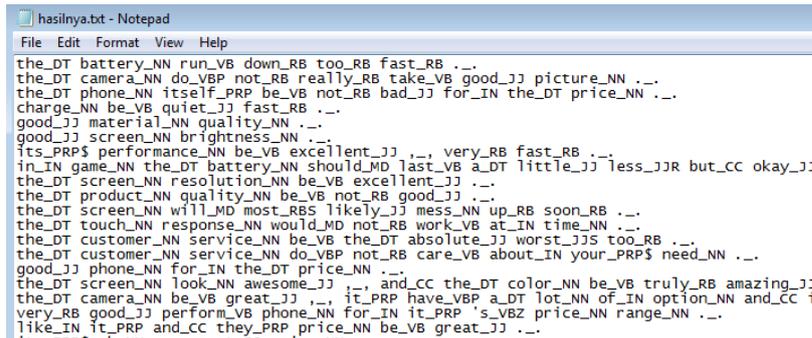
Lemmatization merupakan tahapan untuk mengembalikan kata menjadi kata dasarnya. Tahapan ini dilakukan secara *online* pada *website* <http://textanalysisonline.com/nltk-wordnet-lemmatizer>

3.3. Preprocessing

Proses *preprocessing* yang dilakukan pada tahap ini adalah proses *lowercase*, tokenisasi, *POS tagging* dengan menggunakan Eclipse Mars M1, dan java 1.8.0_71.

a. *Pos Tagging*

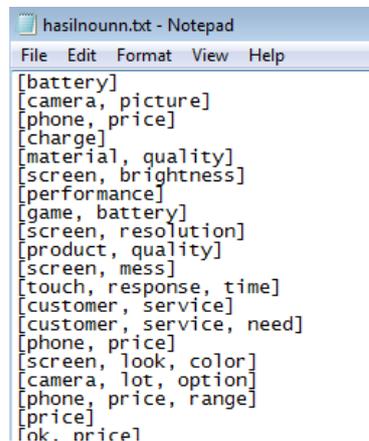
Pos Tagging merupakan tahapan memberikan label perkata kata dari kalimat ulasan. Hasil dari proses *POS Tagging* dapat dilihat pada Gambar 3 dibawah ini.



Gambar 3. Hasil Proses POS Tagging

b. *Noun*

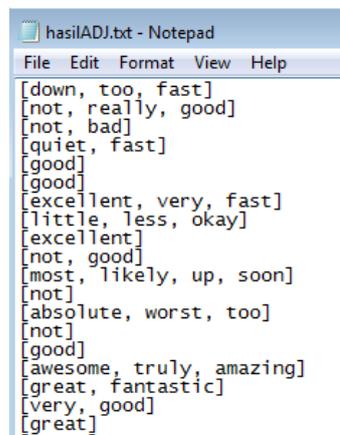
Hasil keluaran dari proses pencarian *noun* ini dapat dilihat pada gambar 4 dibawah ini. Dari hasil *pos tagging* diolah menggunakan javascript untuk memilah dan menemukan *noun* dari kalimat ulasan tersebut.



Gambar 4. Hasil Pencarian *Noun*

c. *Adjective*

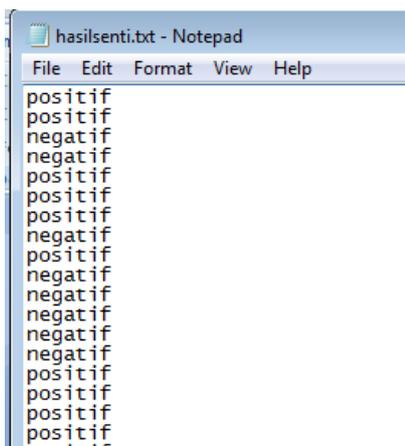
Hasil keluaran dari proses pencarian *adjective* ini dapat dilihat pada Gambar 5 berikut ini. Seperti pencarian *noun*, pencarian *adjective* juga dilakukan menggunakan javascript.



Gambar 5. Hasil *Adjective*

d. Hasil dari Sentimen Analisis

Berdasarkan *adjective* didapat selanjutnya dilakukan sentimen analisis. Hasil tersebut dapat dilihat pada gambar 6 dibawah ini.



Gambar 6. Hasil Sentimen Analisis

3.4. Implementasi RapidMiner

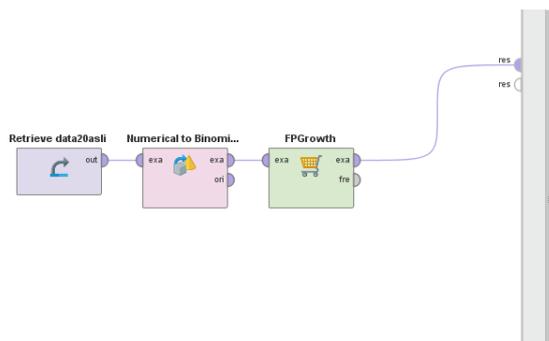
Pengujian dilakukan untuk dapat membuktikan bahwa data yang telah tersedia dapat menghasilkan pola-pola hubungan kombinasi *items* sesuai dengan algoritma *FP-Growth*. Data inputan pada proses *FP-Growth* dibuat menggunakan *Microsoft Excel 2007* berdasarkan *noun* yang dihasilkan pada proses *Get noun*. Data terdiri dari kolom atribut sebagai nama *noun*, setiap *noun* diberi kode *true* untuk kata yang dibicarakan dan *false* untuk kata yang tidak dibicarakan. Gambaran data masukan tersebut dapat dilihat pada Gambar 7 berikut ini.

Pada aplikasi *RapidMiner* yang digunakan terdapat fasilitas *repositories* yang berguna untuk melakukan *add* data masukan yang telah dibuat sebelumnya menggunakan aplikasi *Microsoft Excel 2007*.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	battery	brightness	camera	charge	color	customer	game	look	lot	material	mess	need	ok
2	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
3	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
4	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
5	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
6	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE
7	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
8	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
9	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
10	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
11	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE
12	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
13	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
14	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
15	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE
16	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
17	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
18	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
19	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
20	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
21	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE

Gambar 7. Inputan *FP Growth*

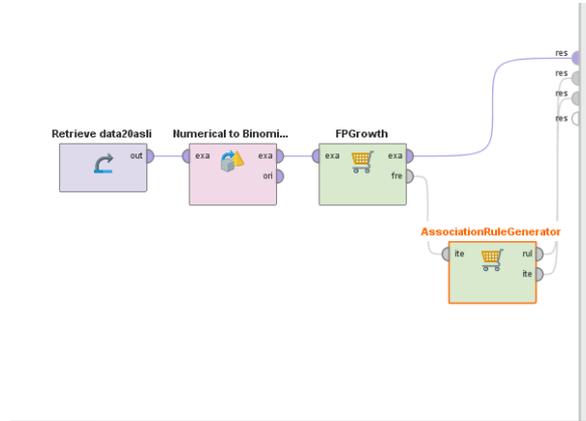
Tahap ini merupakan tahap pemodelan algoritma *FP-Growth*, dengan cara memasukkan operator *FP-Growth* yang terdapat pada jendela *operators* yang kemudian dihubungkan dengan operator *Numerical to Binominal*. Pada pengujian ini ditentukan nilai *minimum support* = 0.1 (10%).



Gambar 8. Penerapan *FP Growth* di *Rapidminer*

a. *Association Rule*

Selanjutnya digunakan operator *Create Association Rules* yang dihubungkan dengan operator *FP-Growth* sebelumnya untuk menghasilkan *rule-rule* dari proses yang dilakukan.



Gambar 9. *Association Rule*

b. Data 420 Interaksi

Pada proses ini dilakukan pengujian terhadap data berjumlah 420 transaksi fitur dengan jumlah atribut sebanyak 366 buah. Pengujian dilakukan dengan melakukan perubahan pada nilai *support* dan *confidence* nya. Hasil dari pengujian ini dapat dilihat pada tabel berikut ini.

Tabel 1. Hasil Pengujian *FP-Growth* 420 Interaksi

No	Support	Confidence	Jumlah rule
1	0,01	0,01	62
2	0,01	0,04	55
3	0,01	0,08	48
4	0,01	0,1	44
5	0,01	0,2	39
6	0,01	0,4	22
7	0,01	0,6	10
8	0,02	0,01	28
9	0,02	0,04	18
10	0,02	0,08	15
11	0,02	0,1	13
12	0,02	0,2	11
13	0,02	0,4	6
14	0,02	0,6	3
15	0,03	0,01	10
16	0,03	0,05	10
17	0,03	0,1	8
18	0,03	0,2	7
19	0,03	0,4	4
20	0,03	0,6	2
21	0,04	0,01	6
22	0,04	0,1	6
23	0,04	0,2	5
24	0,04	0,4	3
25	0,04	0,6	2

Berdasarkan hasil pengujian tersebut dapat disimpulkan bahwa nilai *support*(dukungan) dan *confidence*(kepercayaan) berpengaruh terhadap jumlah *rule* yang dihasilkan. Pemilihan nilai *support* dan *confidence* yang tepat, dapat menghasilkan *rule* yang lebih baik. Semakin besar nilai

support dan nilai *confidence* yang dipilih akan menghasilkan *rule* yang lebih dipercaya. Berikut 6 *rule* yang dihasilkan berdasarkan nilai *support*=4%, dan *confidence*= 10% :

- 1) *If* membicarakan [phone]
Then membicarakan [battery] (confidence: 0.106)
- 2) *If* membicarakan [phone]
Then membicarakan [price] (confidence: 0.237)
- 3) *If* membicarakan [battery]
Then membicarakan [phone] (confidence: 0.309)
- 4) *if* membicarakan [battery]
Then membicarakan [life] (confidence: 0.436)
- 5) *if* membicarakan [price]
Then membicarakan [phone] (confidence: 0.667)
- 6) *if* membicarakan [life]
Then membicarakan [battery] (confidence: 1.000)

c. Hasil *Opinion Words*

Pada Tabel 2 menunjukkan adanya perbedaan hasil antara pencarian manual yang mengandalkan prasangka pembaca dengan pencarian berdasarkan rumus yang berdasarkan nilai polaritas. Hal tersebut dikarenakan adanya kata yang hanya mengandung *adverb* tanpa ada *adjective* seperti *not*, *fast* dan lainnya.

Tabel 2. Hasil *Opinion Word*

No	Jumlah Kalimat	Jumlah <i>opinion words</i> manual		Jumlah <i>opinion words</i> dengan rumusan	
		Positif	Negatif	Positif	Negatif
1	20	13	7	12	8

Pengujian ini selanjutnya dilakukan dengan data lebih banyak sejumlah 420 buah dengan 366 atribut. Hasil dari proses ini dapatkan 201 kata positif dan 219 kata negatif. Oleh karena itu dapat disimpulkan bahwa ulasan yang diberikan pengunjung terhadap Zenfone2 adalah 219 kata negatif dengan fitur yang sering dibicarakan dengan *support* =0,04 adalah phone battery, phone price, dan battery life dan *rule* sebanyak 6 buah dari *confidence*=0,1 .

3.5. Diskusi

Ulasan yang diberikan pembeli pada suatu produk dalam sebuah *e-commerce* tentunya dapat dijadikan sebagai acuan bagi calon pembeli yang berencana akan membeli produk tersebut. Penelitian kali ini berfokus pada pengambilan informasi penting dari kalimat ulasan yang telah diberikan oleh pembeli, lalu melakukan berbagai tahap *preprocessing* guna mendapatkan kata yang sesuai dengan tujuan pengecekan selanjutnya. Lalu dilakukan pengolahan data kata tersebut dengan rapidminer menggunakan algoritma FP-Growth. Data ulasan yang diambil adalah ulasan terkait handphone Asus Zenfone2 dengan kesimpulan yang didapat dari 420 data ada 201 kata positif dan 219 kata negatif.

Berdasarkan hasil pengujian yang telah dilakukan dapat disimpulkan bahwa nilai *support*(dukungan) dan *confidence*(kepercayaan) berpengaruh terhadap jumlah *rule* yang dihasilkan. Pemilihan nilai *support* dan *confidence* yang tepat, dapat menghasilkan *rule* yang lebih baik. Semakin besar nilai *support* dan nilai *confidence* yang dipilih akan menghasilkan *rule* yang lebih dipercaya.

Fp-Growth pernah digunakan dalam review produk berbahasa indonesia (Dhahyani tetya arum, 2015). Penelitian tersebut dilakukan pada ulasan beberapa jenis handpone yaitu nokia lumia, samsung galaxy dan sony xperia. Algoritma yang digunakan sama dengan yang digunakan penulis dalam jurnal kali ini, namun berbeda dengan *preprocessing* , dan proses penelitiannya. Dari penelitian ini didapat nilai rata-rata *accuracy* untuk proses orientasi opini dari data set Nokia Lumia, Samsung Galaxy, dan Sony Xperia adalah 71,9%. Orientasi opini dipengaruhi oleh kelengkapan dan ketepatan definisi synset pada SentiwordNet. Semakin banyak kata dalam suatu opini yang tidak tercantum pada Sentiwordnet mengakibatkan banyak kata tidak dapat dibobotkan dengan tepat, sehingga hasil klasifikasi orientasi

opini menjadi tidak tepat. Penelitian terkait Ulasan Produk lainnya juga pernah dilakukan sebelumnya namun menggunakan Metode Naive Bayes (Billy Gunawan, et al, 2018).

Secara keseluruhan, penelitian ini dapat memberikan gambaran terkait pemanfaatan data ulasan sebagai media informasi dalam mengambil keputusan bagi masyarakat dan perusahaan terkait. Dengan pengolahan data yang tepat maka akan membantu masyarakat dalam memilih produk yang tepat tanpa melakukan pengecekan ulasan satu-persatu. Begitu juga bagi perusahaan terkait, data informasi hasil pengolahan tersebut dapat digunakan sebagai evaluasi perusahaan.

4. KESIMPULAN

Proses *preprocessing* yang dilakukan meliputi *lowercase*, tokenisasi, dan POS Tagging, dan penerapan *opinion words* dapat diterapkan dengan *tools* yang penulis buat. Algoritma *FP-Growth* dapat diterapkan untuk menentukan fitur dari sebuah handphone yang sering dibicarakan oleh pengguna media sosial. Informasi mengenai fitur tersebut dapat dimanfaatkan oleh para calon pembeli sebagai bahan pertimbangan dalam membeli produk, sedangkan untuk pihak pemilik produk dapat dijadikan sebagai bahan evaluasi. algoritma *FP Growth* melakukan pemindaian kumpulan data transaksi hanya dilakukan dua kali, sehingga tidak dilakukan pemindaian data transaksi secara berulang-ulang. Dan *Opinion Words* dapat digunakan untuk menentukan apakah ulasan yang diberikan pengunjung terhadap suatu produk bernilai positif ataupun negatif. Hasil ini dapat dijadikan juga oleh pengunjung sebagai bahan pertimbangan dalam pemilihan produk, serta dapat dijadikan sebagai bahan evaluasi oleh pihak pemilik produk.

Penelitian ini dilakukan terkait data ulasan produk asus zenfone2 yang telah dilakukan berdasarkan data sejumlah 420 buah dengan 366 atribut. Hasil dari proses ini dapatkan 201 kata positif dan 219 kata negatif. Oleh karena itu dapat disimpulkan bahwa ulasan yang diberikan pengunjung terhadap Zenfone2 adalah 219 kata negatif dengan fitur yang sering dibicarakan dengan *support* =0,04 adalah phone battery, phone price, dan battery life dan *rule* sebanyak 6 buah dari *confidence*=0,1.

DAFTAR PUSTAKA

- BrightLocal. (2020). Local Consumer Review Survey 2020. Available Online at: <https://www.brightlocal.com/research/local-consumer-review-survey-2020>
- Erwin. (2009). Analisis Market Basket dengan Algoritma Apriori dan FP-Growth. *Jurnal Generic* , 4. Association Rule Metode Apriori. *Informasi dan Teknologi Ilmiah (INTI)* , III.
- Haque, T. U., Saber, N. N., & Shah, F. M. (2018). Sentiment analysis on large scale Amazon product reviews. 2018 IEEE International Conference on Innovative Research and Development (ICIRD). doi:10.1109/icird.2018.8376299
- Liu, Bing. (2020). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press
- Kurniawan, B., Effendi, S., & Salim Sitompul, O. (2012). Klasifikasi Konten Berita dengan Metode Text Mining. *Jurnal Dunia Teknologi Informasi* , 1, 14-19.
- Litbang. (2019). *APJII: Jumlah Pengguna Internet di Indonesia Tembus 171 Juta Jiwa*. Litbang Kemendagri. <https://litbang.kemendagri.go.id/website/apjii-jumlah-pengguna-internetdiindonesia-tembus-171-juta-jiwa/>
- Milatina, Syukur, A., & Supriyanto, C. (2012). Pengaruh Text Preprocessing pada Clustering Dokumen Teks Berbahasa Indonesia. *Jurnal Teknologi Informasi* , 8.
- Radovanovic, M., & Ivanovic, M. (2008). *Text Mining: Approaches and Application*.
- Ririanti. (2014). Implementasi Algoritma FP-Growth pada Aplikasi Prediksi Persediaan Sepeda Motor (Studi Kasus PT.Pilar Deli Labumas). *Pelita Informatika Budi Darma* (VI).
- Vanaja, S., Belwal, M. (2018). Aspect-Level Sentiment Analysis on E-Commerce Data. Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018). doi:10.1109/ICIRCA.2018.8597286

Halaman Ini Dikosongkan