ANALISIS OPINI PADA MEDIA SOSIAL FACEBOOK DENGAN MENGUNAKAN TEXT MINNING TERHADAP PRODUK HIJAB (STUDI KASUS : ZOYA DAN EL ZATTA)

Agus Heryanto, S.T., M.Kom

Dosen Prodi Manajemen Informatika

Politeknik LP3I Bandung

E-Mail : agusvalentino46@gmail.com

Abstrak : Analisis opini atau yang sering disebut juga dengan istilah sentimen analysis merupakan analisis yang bertujuan untuk melihat opini masyarakat atau kelompok mengenai entitas tertentu. Sentimen analysis sendiri sering digunakan untuk menilai sebuah produk, apakah memberikan review yang baik dan disukai masyarakat ataupun sebaliknya. Penelitian ini dilakukan untuk mengklasifikasikan opini atau sentimen masyarakat terhadap produk hijab. Hijab sekarang menjadi trend dikalangan remaja muslim, dimana sebagian besar remaja muslim merupakan para pengguna media sosial terutama *facebook*. Sehingga pandangan para pengguna hijab ini akan tertuang dalam sebuah persepsi, baik itu positif maupun negatif. Oleh karena itu, penelitian ini mengambil opini-opini masyarakat terhadap produk hijab melalui *facebook* dengan dua *fanspage* dari dua produk hijab yaitu zoya dan el-zatta. Metode penghitungan dan pengklasifikasiian yang dipakai adalah metode *naive bayes*. Hasil penelitian menunjukkan bahwa sentiment positif terbentuk paling banyak yaitu 925 komentar pada EL-zatta, dan 882 komentar pada Zoya. Jika dibandingkan dengan sentimen negatif yang hanya 71 komentar pada El-zatta dan 125 komentar pada Zoya. Jadi secara keseluruhan, persepsi masyarakat mengenai dua produk hijab ini yang tertuang pada komentar di *fanspage facebook* memberikan penilaian yang positif.

Kata kunci : Analisis Opini, *Naive Bayes*, *Facebook*, Produk Hijab

1. Pendahuluan

Alat bertukar informasi dan komunikasi modern saat ini yang berkembang pesat di masyarakat adalah media sosial. Bahkan sosial media ini bagian hidup masyarakat menengah ke atas. Sehingga media sosial, menjadi wadah dalam berkomunikasi yang menyediakan ruang interaksi tanpa batas tempat dan waktu. Oleh karena itu, informasi yang didapat pun akan selalu terbaru dan berasal dari manapun yang memang itu dibutuhkan oleh masyarakat.

Salah satu media sosial yang berkembang dan diminati saat ini adalah *facebook*. Bersama dengan berbagai media sosial yang lain, *facebook* menjadi media sosial yang banyak di akses termasuk oleh masyarakat indonesia. Sehingga tak heran jika berdasarkan beberapa survey, menunjukan bahwa indonesia ada dipreingkat ke-4 didunia sebagai pengguna paling banyak (Okezone dikutip dari Kominfo 2014). Dan ternyata remaja dan wanita adalah kalangan yang paling banyak dan paling aktif dalam menggunakan *facebook*.

Disisi lain, remaja perempuan diindonesia yang sebagian besarnya beragama islam sekarang ini sudah banyak yang berhijrah. Ini ditandai dengan banyaknya perempuan muslimah yang berhijab. Bahkan sekarang hijab sudah menjadi tren masa kini yang desain dan tampilannya banyak diminati oleh para remaja. Dimana sebagian besar remaja itu adalah pengguna media sosial termasuk *facebook*.

Beberapa produsen hijab diindonesia memanfaatkan sosial media untuk memperkenalkan dan memasarkan produknya kepada masyarakat. Dari banyaknya produk hijab yang juga sekarang ini menjamur di masyarakat, dalam penelitian ini akan diambil dua produk hijab yaitu : Zoya dan El-Zatta.

Banyaknya pertukaran informasi di media sosial *facebook* baik itu postingan berupa status, komentar ataupun percakapan dan yang lainnya, cenderung bersifat jujur dan spontan diungkapkan. Hal tersebut, dapat dimanfaatkan untuk mengetahui tingkat popularitas suatu produk, kepuasan pelanggan dari sebuah produk, dan apa yang pelanggan harapkan dari sebuah produk. Itu semua menunjukkan persepsi masyarakat dari sebuah produk. Sehingga ini akan menjadi bahan yang sangat penting untuk pengembangan bisnis kedepannya bagi sebuah usaha.

Untuk itu perlu suatu analisis opini ataupun opini masyarakat terhadap produk hijab, yang dalam penelitian ini produk hijabnya adalah Zoya dan El-zatta. Analisis opini yang dihasilkan dari persepsi masyarakat, bisa opini positif ataupun negatif

2. Pemahaman Konsep

2.1. Analisis Sentimen

Analisis sentimen juga dapat dikatakan sebagai opinion mining. Analisis sentimen dapat digunakan dalam berbagai kemungkinan domain, dari produk konsumen, jasa kesehatan, jasa keuangan, peristiwa sosial dan politik pada pemilu. Kecenderungan penelitian tentang analisis sentimen berfokus pada pendapat yang menyatakan atau menyiratkan suatu sentimen positif atau negatif. Pendapat mewakili hampir semua aktivitas manusia, karena pendapat dapat mempengaruhi terhadap perilaku seseorang. Setiap kali kita perlu membuat keputusan, kita ingin tahu pendapat orang lain. Dalam dunia nyata, bisnis dan organisasi selalu ingin melihat opini publik tentang suatu produk atau jasa (Liu, 2012).

Analisis sentimen pada suatu kalimat menggambarkan bagian pertimbangan penilaian terhadap entitas atau kejadian tertentu. Entitas adalah (Liu, 2012) produk, layanan, topik, isu, orang, organisasi, atau peristiwa yang menjadi objek target pada kalimat sentimen. Suatu hal yang penting, untuk memahami permasalahan dalam membedakan kalimat sentimen.

2.2. Text Mining

Text mining, pada proses mengambil informasi dari teks. Informasi biasanya diperoleh melalui peramalan pola dan kecenderungan pembelajaran pola statistik. *Text mining* yaitu parsing, bersama dengan penambahan beberapa fitur linguistik turunan dan penghilangan beberapa diantaranya, dan penyisipan subsequent ke dalam *database*, menentukan pola dalam data terstruktur, dan akhirnya mengevaluasi dan menginterpretasi output, *text mining* biasanya mengacu ke beberapa kombinasi relevansi, kebaruan, dan *interestingness*. Proses *text mining* yang khas meliputi kategorisasi teks, *text clustering*, ekstraksi konsep/entitas, produksi taksonomi granular, sentiment analysis, penyimpulan dokumen, dan pemodelan relasi entitas yaitu, pembelajaran hubungan antara entitas (Bridge, 2011).

Pendekatan manual *text mining* secara intensif dalam laboratorium pertama muncul pada pertengahan 1980-an, namun kemajuan teknologi telah memungkinkan ranah tersebut untuk berkembang selama dekade terakhir. *Text mining* adalah bidang interdisipliner yang mengacu pada pencarian informasi, pertambahan data, pembelajaran mesin, statistik, dan komputasi linguistik. Dikarenakan kebanyakan informasi (perkiraan umum mengatakan lebih dari 80%) saat ini disimpan sebagai teks, *text mining* diyakini memiliki potensi nilai komersial tinggi (Bridge, 2011).

Text mining merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar (Fieldman, dkk. 2007).

Langkah-langkah yang dilakukan dalam *text mining* adalah sebagai berikut :

1. *Text Preprocessing*

Tindakan yang dilakukan pada tahap ini adalah *toLowerCase*, yaitu mengubah semua karakter huruf menjadi huruf kecil, dan *Tokenizing* yaitu proses penguraian deskripsi yang semula berupa kalimat – kalimat menjadi kata-kata dan menghilangkan delimiter-delimiter seperti tanda titik(.), koma(,), spasi dan karakter angka yang ada pada kata tersebut (Wiess, dkk. 2005).

2. *Feature Selection*

Pada tahap ini tindakan yang dilakukan adalah menghilangkan *stopword* (*stopword removal*) dan stemming terhadap kata yang berimbuhan (Berry, 2010). *Stopword* adalah kosakata yang bukan merupakan ciri (kata unik) dari suatu dokumen. Misalnya “di”, “oleh”, “pada”, “sebuah”, “karena” dan lain sebagainya. Stemming adalah proses pemetaan dan penguraian berbagai bentuk (variants) dari suatu kata menjadi bentuk kata dasarnya (stem). Tujuan dari proses stemming adalah menghilangkan imbuhan-imbuhan baik itu berupa prefiks, sufiks, maupun konfiks yang ada pada setiap kata

2.3 Algoritma *Naïve Bayes Classifier*

Sebuah *bayes classifier* adalah *classifier* probabilistik sederhana berdasarkan penerapan teorema Bayes (dari statistik Bayesian) dengan asumsi independen (naif) yang kuat. Sebuah istilah yang lebih deskriptif untuk model probabilitas yang digaris bawahi adalah "model fitur independen".

Dalam terminologi sederhana, sebuah NBC mengasumsikan bahwa kehadiran (atau ketiadaan) fitur tertentu dari suatu kelas tidak berhubungan dengan kehadiran (atau ketiadaan) fitur lainnya. Sebagai contoh, buah mungkin dianggap apel jika merah, bulat, dan berdiameter sekitar 4 inchi. Bahkan jika fitur ini bergantung satu sama lain atau atas keberadaan fitur lain,. Sebuah NBC menganggap bahwa seluruh sifat-sifat berkontribusi mandiri untuk probabilitas bahwa buah ini adalah apel.

Tergantung pada situasi yang tepat dari model probabilitas, NBC dapat dilatih sangat efisien dalam *supervised learning*. Dalam aplikasi praktis, parameter estimasi untuk model NBC menggunakan metode *likelihood* maksimum, dengan kata lain, seseorang dapat bekerja dengan model Naïve Bayes tanpa mempercayai probabilitas Bayesian atau menggunakan metode Bayesian lainnya.

Dibalik desain naifnya dan asumsi yang tampaknya terlalu disederhanakan, NBC telah bekerja cukup baik dalam banyak situasi dunia nyata yang kompleks. Pada tahun 2004, analisis masalah klasifikasi Bayesian telah menunjukkan bahwa ada beberapa alasan teoritis untuk keberhasilan yang tampaknya tidak masuk akal dari NBC (Zhang, H., 2004). Selain itu, perbandingan yang komprehensif dengan metode klasifikasi lainnya pada tahun 2006 menunjukkan bahwa klasifikasi Bayes mengungguli pendekatan terbaru, seperti *boosted tree* atau *random forest* (Caruana, R. & Niculescu-Mizil, A, 2006).

Sebuah keuntungan dari NBC adalah bahwa ia memerlukan sejumlah kecil data pelatihan untuk mengestimasi parameter (rata-rata dan varian dari variabel) yang diperlukan untuk klasifikasi. Karena variabel diasumsikan independen, hanya varian dari variabel-variabel untuk setiap kelas yang perlu ditentukan dan bukan keseluruhan *covariance matrix*.

3. Metodologi Penelitian

Metododologi yang digunakan dalam penelitian ini yaitu *Naive Bayes Classifier*.

Langkah yang dilakukan untuk melakukan penelitian ini adalah dengan melakukan identifikasi dan analisis masalah terlebih dahulu. Setelah itu dilanjutkan dengan tahapan pengumpulan atau pengambilan data dari media sosial *facebook*, yang akan dibutuhkan untuk melkaukan penelitian, kemudian data tersebut diolah dengan menggunakan algoritma yang telah ditetapkan sehingga akan didapat hasil yang sesuai dengan yang dibutuhkan.

Secara garis besar langkah penelitian disajikan dengan menggunakan diagram yang ditunjukan pada gambar 3.2. Namun untuk lebih jelas akan diuraikan terlebih dahulu langkah -langkah penelitian yang akan dilakukan :

1. Identifikasi dan analisa kebutuhan data

Tahap identifikasi dan analaisis kebutuhan data dilakukan untuk mengetahui kebutuhan informasi yang dubutuhkan, sehingga hasil dari penelitian ini sesuai dengan uyang diharapkan.

2. Studi literature

Pada tahap Studi Literature peneliti mencari dan kemudian menganalisa jurnal – jurnal yang berhubungan dengan penelitian sehingga dapat menjadi pembanding ataupun menjadu sumber pustakan bagi penelitian.

Setelah jurnal – jurnal yang berkaitan dibaca dengan seksama, maka dipilih 10 jurnal yang memiliki hubungan dan keterkaitan dengan penelitian ini yaitu literatur yang diambil adalah tentang TF – IDF untuk pembobotan term dari data yang di analisis dan algoritma naive bayes untuk penghitungan sentimen analisisnya.

3. Data selection

Tahap ini dilakukan untuk mencari data yang akan diolah. Data yang masih kotor tersebut harus dianalisis terlebih dahulu agar pada saat proses pengolahan data, sumber data yang digunakan lebih efektif dan efisien. Selain itu pula dengan adanya analisis sumber data, dapat diketahui mana data yang tidak dapat digunakan dan dapat pula mengetahui karakteristik data yang akan diolah.

Sumber data yang digunakan pada penelitian ini adalah data yang diambil dari *facebook*. Data tersebut merupakan data dari beberapa fanspage yang telah ditentukan sesuai dengan kasus yang diambil. Dan yang lebih spesifik adalah data komentar dari para pengguna *facebook*. Kata – kata dari komentar pun yang diambil adalah data berbahasa indonesia. Pengambilan data dari fanspage *facebook* ini menggunakan aplikasi yang telah disediakan oleh *facebook* sendiri yaitu *netvizz*.

4. Pre – processing

Tahapan *preprocessing* terdiri atas *tokenizer*, penghilangan stopwords, stemming dan pembobotan. *Tokenizer* adalah proses pemecahan dokumen teks menjadi beberapa token atau kata, pembatasan biasanya dilakukan berdasarkan spasi atau tanda baca. Tahap selanjutnya adalah proses penghapusan kata-kata yang sesuai pada daftar *stopwords*. *Stopwords* merupakan daftar kata-kata yang dianggap tidak merepresentasikan konten dari suatu dokumen teks. Kata-kata yang tersisa dari tahapan ini kemudian dilakukan proses stemming, yaitu merubah kata menjadi bentuk dasarnya (Pabrin, 2013).

5. Data mining

Klasifikasi dalam sentimen analisis dapat dilakukan dengan berbagai macam algoritma, salah satunya adalah dengan metode *Naïve bayes*.

Naive bayes merupakan metode klasifikasi dalam text mining berdasarkan probabilitas dan teorema bayes. Adapun asumsi yang harus dipenuhi yaitu adanya independensi dalam variable bebas (Destuardi, 2009). Maksudnya adalah kemunculan sebuah kata tidak mempengaruhi kemunculan kata lainnya, ataupun sebaliknya.

Analisis *bayesian* menggali dua buah informasi tentang parameter suatu model statistik, yang mana bila digabungkan informasi ini akan membentuk informasi posterior. Peluang dari distribusi posterior dari parameter yang tidak diketahui diperoleh dari data dan juga distribusi priornya. Metode bayesian menyelidiki bentuk hubungan antara distribusi prior dan fungsi *likelihood*-nya. Algoritma *Naive*

Bayes Classifier dengan asumsi semua atribut independen dapat dituliskan sebagai berikut (Eirene, 2014):

$$\begin{aligned} P(X_1, X_2, \dots, X_n|Y) &= P(X_1|X_2, \dots, X_n, C)P(X_2, \dots, X_n|Y) \\ &= P(X_1|C)P(X_2, \dots, X_n|Y) \\ &= P(X_1|C)P(X_2|C) \dots P(X_n|C) \end{aligned}$$

Formula (2) dapat diterapkan pada klasifikasi dokumen teks. Misalkan diketahui bahwa peluang term ke- i pada dokumen D terjadi di suatu kelas A adalah $P(t_i|A)$ Maka peluang dokumen tersebut terjadi di kelas A adalah $P(D|A) = \pi_t P(t_i|A)$ Apabila kelas terdiri dari kelas positif (A_{pos}) dan kelas negatif (A_{neg}), maka peluang dokumen D masuk kelas positif adalah

$$P(A_{pos}|D) = \frac{P(A_{pos})P(D|A_{pos})}{P(D)} = \frac{P(A_{pos})\pi_t P(Tt|A_{pos})}{P(D)}$$

Demikian juga untuk peluang D masuk dalam dokumen negatif

$$P(A_{neg}|D) = \frac{P(A_{neg})P(D|A_{neg})}{P(D)} = \frac{P(A_{neg})\pi_t P(Tt|A_{neg})}{P(D)}$$

Rasio dari keduanya diperoleh:

$$\frac{P(A_{pos}|D)}{P(A_{neg}|D)} = \frac{P(A_{pos})}{P(A_{neg})} + \prod_t \frac{P(Tt|A_{pos})}{P(Tt|A_{neg})}$$

dengan mengambil logaritma di setiap ratio diatas diperoleh

$$\ln \frac{P(A_{pos}|D)}{P(A_{neg}|D)} = \ln \frac{P(A_{pos})}{P(A_{neg})} + \sum_t \ln \frac{P(Tt|A_{pos})}{P(Tt|A_{neg})}$$

Dari persamaan (3), maka klasifikasi dokumen akan terbentuk. Dokumen dikategorikan positif jika $P(A_{pos}|D) > P(A_{neg}|D)$ atau $\ln \frac{P(A_{neg}|D)}{P(A_{pos}|D)} > 0$, demikian pula sebaliknya.

Sedangkan dokumen akan masuk kelas netral jika $\ln \frac{P(A_{neg}|D)}{P(A_{pos}|D)} = 0$,

Gambar 2 Metodologi Penelitian

5. Hasil dan Pembahasan

Total data yang dilakukan pengamatan sebanyak 2000 dokumen yang terdiri dari 1000 dokumen dari fanspage Zoya Lovers, dan 1000 dokumen dari fanspage El-zatta Hijab ke semua dokumen tersebut dapat dilihat pada lampiran B dan C. Setelah melalui beberapa tahapan pada preprocessing, dan kemudian dilanjutkan dengan penghitungan bobot serta pengklasifikasian menggunakan metode naive bayes. Maka diperoleh berbagai hasil seperti dibawah ini.

Gambar 3 Sentimen Analisis Fanspage Zoya

Pada fanspage Zoya Lovers, dari 1000 dokumen yang berupa komentar dari para pengguna facebook. Dapat diklasifikasikan kedalam sentimen positif sebanyak 882 komentar atau 87,1%, sedangkan sentimen negatif sebanyak 125 komentar atau 12,3 %, serta terdapat 5 komentar netral atau 0,5%. Sehingga dapat terlihat bahwa respon atau pandangan masyarakat terhadap produk ataupun layanan dari zoya tergolong sangat baik.

Dari kata yang sentimen negatif yang dikumpulkan pada fanspage Zoya Lovers, kemudian dilakukan pencarian kata yang paling banyak muncul dengan menggunakan *wordcloud* yang diakses pada situs worditout.com. Ternyata kata yang paling banyak muncul adalah harga sebanyak 52 kata, kemudian berturut – turut selanjutnya ada kata beli (24), turun (23), diskon (15) dan promo (12).

Gambar 4 Wordcloud Fanspage Zoya

Sedangkan pada fanspage El-zatta Hijab, dari 1000 dokumen yang berupa komentar dari para pengguna facebook. Dapat diklasifikasikan kedalam sentimen positif sebanyak 925 komentar atau 92,5%, sedangkan sentimen negatif sebanyak 71 komentar atau 7,1 %, serta terdapat 4 komentar netral atau 0,4%. Sehingga dapat terlihat bahwa respon atau pandangan masyarakat terhadap produk ataupun layanan dari El-zatta tergolong sangat baik, bahkan jika dibandingkan dengan Zoya responnya lebih baik

Gambar 5 Sentimen Analisis Fanspage Elzatta

Lalu dari kata yang sentimen negatif yang dikumpulkan pada fanspage El-zatta Hijab, kemudian dilakukan pencarian kata yang paling banyak muncul dengan menggunakan *wordcloud* yang juga diakses melalui situs worditout.com. Ternyata kata yang paling banyak muncul adalah web sebanyak 14 kata, kemudian berturut – turut selanjutnya ada kata agen (11), buka (11), beli (9) dan diskon (9).

Gambar 6 Wordcloud Elzatta

6. Kesimpulan

Berdasarkan uraian di atas, dapat disimpulkan hal-hal sebagai berikut:

1. Metoda Naïve Bayes Classifier dapat diterapkan sebagai metoda untuk melakukan klasifikasi sentiment analysis.
2. Dari berbagai analisis dapat diberikan hasil bahwa respon atau pandangan masyarakat terhadap hijab tergolong sangat baik. Untuk pembagiannya sendiri, pada

fanspage Zoya Lovers dengan klasifikasi komentar yang masuk kedalam sentimen positif adalah sebanyak 882 dokumen, sentimen negatif sebanyak 125 dokumen dan kelas netral sebanyak 5. Sedangkan untuk fanspage El-zatta Hijab dengan klasifikasi komentar pada sentimen positif sebanyak 925 dokumen, sentimen negatif sebanyak 71 dan kelas netral sebanyak 4 komentar.

3. Kemudian berdasarkan wordcloud, dapat disimpulkan juga kata pada sentimen negatif yang paling sering muncul adalah kata harga pada fanspage Zoya Lovers, dan kata web pada fanspage El-zatta Hijab.

7. Saran

Saran pengembangan penelitian ini yaitu agar memperoleh hasil yang lebih maksimal data training yang digunakan sebaiknya lebih banyak agar tingkat akurasi pengklasifikasian sentimennya lebih akurat. Juga untuk data testing sumber data yang diambil sebaiknya dengan rentang waktu yang lebih panjang dan juga data yang lebih banyak pula. Untuk studi kasus hijab sendiri, karena sekarang banyak produk hijab, sebaiknya studi kasus nya disarankan lebih banyak agar lebih terlihat perbedaannya, dan menghasilkan hasil yang maksimal.

8. Acknowledgment

Terima kasih disampaikan kepada rekan - rekan dosen Manajemen Informatika LP3I Bandung yang telah membantu dalam proses pengerjaan jurnal ini. Dan kepada Dr. Djaya Kusuma Wijaya, ST. M.Kom yang telah membimbing untuk pengerjaan jurnal ini.

Daftar Pustaka:

- [1] Adriani, M., Asian, J., Nazief, B., Tahaghoghi, S.M.M., Williams, H.E. 2007. *Stemming Indonesian : A Confix - Stripping Approach. Transaction on Asian Language Information Processing. Vol. 6, No. 4, Article 13. Association for Computing Machinery, New York*
- [2] Apriliyanti, Anindiya. 2015. *Sentiment Analysis dengan Naive Bayes untuk Melihat Persepsi Masyarakat Terhadap Batik pada Jejaring Sosial Twitter*. Departemen Statistika FMIPA, Universitas, Bandung
- [3] Berry, M.W. & Kogan, J. 2010. *Text Mining Application and Theory*. WILEY : United Kingdom
- [4] Caruana, R. & Niculescu-Mizil, A. 2006. *An Empirical Comparison of Supervised Learning Algorithms. Proceedings of The 23rd International Conference on Machine Learning*.
- [5] Destuardi dan Sumpeno, Surya. 2009. *Klasifikasi Emosi Teks Bahasa Indonesia Menggunakan Metode Naive Bayes*. Institut Teknologi Sepuluh Nopember. Surabaya
- [6] Dragut, E., Fang, F., Sistla, P., Yu, S. & Meng, W. 2009. *Stop Word and Related Problems In Web Interface Integration*.
- [7] Feldman, R & Snager, J. 2007. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press, New York.

- [8] Imam Adiyana. 2015. *Implementasi Text Mining Pada Mesin Pencarian Twitter Untuk menganalisis topik – Topik Terkait “Kpk Dan Jokowi”*. Universitas Islam Indonesia.
- [9] Lama, Prabin. 2013. *Clustering System Based on Text Mining Using The K-Means Algorithm*. Thesis : Turku Universitas of Applied Sciences, Southwest Finland
- [10] Larose, & Daneil T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. USA : John Wiley and Sons
- [11] Liu, Bing. 2012. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publisher.
- [12] Mitra, Mandar. 2014. *Introduction to Text Mining*. Indian Statistical Institute.
- [13] Nur,Y., Santika, D. D. 2011. *Analisis Sentimen Pada Dokumen Berbahasa Indonesia Dengan Pendekatan Support Vector Machine*. *Konferensi Nasional Sistem dan Informatika 2011*; Bali.
- [14] Pramudiono, I. 2007, *Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*
- [15] Pudjiantoro, Tachbir Hendro, dkk. 2014. *Perangkat Lunak Pengambilan Pengetahuan Akuntansi Di dalam Big Data*. Jurusan Informatika, UNJANI, Cimahi
- [16] Retnawiyat, Eka dkk. 2015. *Analisis Sentimen Pada Data Twitter dengan Menggunakan Text Mining terhadap Suatu Produk*. Ilmu Komputer
- [17] Rima Aditya, Bayu. 2015. *Pengguna Web Crawler Untuk Menghimpun Tweets Dengan Metode Pre-Processing Text Mining*. Fakultas Ilmu Terapan, Universitas Telkom Bandung.
- [18] Rizal Himawan Sucipto. *Analisis Jaringan Teks Berdasarkan Social Network Analysis Dan Text Mining Untuk Mengetahui Persepsi Kualitas Merek Pada Konten Percakapan Di Media Sosial Twitter (Studi Pada Pt. Indosat Tbk. Dan Pt Telkomsel)*. Fakultas Ekonomi dan Bisnis, Universitas Telkom.
- [19] Siti Mujilahwati. 2016. *Pre-Processing Text Mining Pada Data Twitter*. Fakultas Teknik, Universitas Islam Lamongan
- [20] Sihombing, Eirene. 2014. *Penerapan Analisis Sentimen dengan Metode Naive Bayes pada Klasifikasi Data Teks*. Universitas Padjadjaran. Bandung
- [21] Weiss, S.M., Indurkha, N., Zhang, T., Damerau, F.J 2005. *Text Mining : Predictive Methods to Analyzing Unstructured Information*. Springer : New York.
- [22] Yosef Ganisaputra dan Robby Tan. 2015. *Pembuatan Aplikasi Datamining Facebook dan Twitter dengan Naïve Bayes Classifier*. Fakultas Teknologi Informasi, Universitas Kristen Maranatha, Bandung
- [23] Zhang, H. 2004. *The Optimality of Naive Bayes*.