

SURVEI: TOPIK DAN TREN ANALISIS SENTIMEN PADA MEDIA ONLINE

Sumarni Adi¹⁾, Murni Wulandari²⁾, Anis Kemala Mardiana³⁾, Ahmad Muzakki⁴⁾

^{1), 2), 3), 4)} Fakultas Ilmu Komputer, Program Studi Informatika, Universitas AMIKOM Yogyakarta
Jl. Ring Road Utar, Condong Catur, Depok, Sleman, Yogyakarta 55281

Email : sumarni.a@amikom.ac.id¹⁾, murni.wulandari@students.amikom.ac.id²⁾, anis.kemala@gmail.com³⁾,
Ahmad.muzakki@students.amikom.ac.id⁴⁾

Abstract

In general there are two types of textual information in online media that is facts and opinions. Facts are objective statements about the world's entities and events while opinions are subjective statements that reflect people's sentiments or perceptions of the entities or events of the world. Because of the rich online media resources of opinion such as social media, discussion forums, review sites, blogs and newspapers available in digital form, most of the current research focuses on the field of sentiment analysis. The research undertaken is intended to develop a system that can identify and classify opinions or sentiments in order to gain public opinion about the product, image or service, in order not to need to conduct conventional surveys and focus groups that are old and expensive.

Accurate methods for predicting sentiment can allow us to take opinions from the internet and predict online customer choice, which could prove valuable for economic or marketing research. To date, there are several different problems that dominate this research community, namely sentiment classification, feature-based classification and handling negations. This paper presents a survey that includes techniques and methods in analysis and the challenges of sentiment that appear in online media.

Keywords: *opinion, sentiment, classification, machine learning*

Abstrak

Secara umum terdapat dua tipe informasi tekstual di media online yaitu fakta dan opini. Fakta adalah pernyataan objektif mengenai entitas dan kejadian di dunia sedangkan opini adalah pernyataan subjektif yang merefleksikan sentimen atau persepsi orang mengenai entitas ataupun kejadian di dunia. Karena banyaknya sumber daya media online kaya akan opini seperti sosial media, forum diskusi, review situs, blog dan surat kabar yang tersedia dalam bentuk digital, sebagian besar penelitian saat ini berfokus pada bidang analisis sentimen. Penelitian yang dilakukan dimaksudkan untuk

mengembangkan sebuah sistem yang dapat mengidentifikasi dan mengklasifikasikan opini atau sentimen agar dapat memperoleh opini publik mengenai produk, citra atau layanan, agar tidak perlu melakukan survei konvensional dan fokus group yang lama dan mahal biayanya.

Metode yang akurat untuk memprediksi sentimen dapat memungkinkan kita untuk mengambil opini dari internet dan memprediksi pilihan pelanggan online, yang bisa terbukti berharga untuk penelitian ekonomi atau pemasaran. Hingga saat ini, ada beberapa masalah yang berbeda yang mendominasi komunitas penelitian ini, yaitu klasifikasi sentimen, klasifikasi berbasis fitur dan penanganan negasi. Karya ilmiah ini menyajikan sebuah survei yang mencakup teknik dan metode dalam analisis dan tantangan sentimen yang muncul di media online.

Kata Kunci : *opini, sentimen, klasifikasi, pembelajaran mesin*

1. Pendahuluan

Analisis sentimen disebut juga dengan *opinion mining* (penambangan opini) merupakan proses untuk mengekstrak suatu opini atau pendapat dari dokumen untuk topik tertentu [1]. Analisis sentimen adalah jenis pengolahan bahasa alami untuk melacak suasana hati masyarakat tentang produk atau topik tertentu. Analisis sentimen, yang juga disebut *opinion mining*, melibatkan dalam membangun sebuah sistem untuk mengumpulkan dan memeriksa opini tentang produk atau layanan yang dibuat dalam postingan web, blog, atau komentar pada sosial media. Analisis sentimen bermanfaat dalam banyak hal. Misalnya, dalam pemasaran, hal itu membantu keberhasilan iklan atau peluncuran produk baru, menentukan versi produk atau layanan mana yang populer dan bahkan mengidentifikasi jenis demografi yang menyukai atau tidak menyukai fitur tertentu.

Ada beberapa tantangan dalam analisis sentimen. Yang pertama adalah opini yang dianggap positif dalam satu situasi dan dapat dianggap negatif dalam situasi lain. Tantangan kedua adalah bahwa orang tidak selalu

mengungkapkan pendapat dengan cara yang sama. Sebagian besar pemrosesan teks konvensional bergantung pada fakta bahwa perbedaan kecil antara dua teks tidak terlalu banyak artinya. Namun, dalam analisis Sentimen, "gambarnya bagus" sangat berbeda dari "gambarnya tidak bagus". Orang bisa menjadi kontradiktif dalam pernyataan tersebut. Sebagian besar opini akan memiliki komentar positif dan negatif, yang agak mudah dikelola dengan menganalisis kalimat satu per satu. Namun, di media yang lebih informal seperti twitter atau blog, semakin besar kemungkinan orang menggabungkan pendapat berbeda dalam kalimat yang sama yang mudah dipahami oleh manusia, namun lebih sulit bagi komputer untuk mengurai. Terkadang bahkan orang lain pun sulit memahami apa yang dipikirkan seseorang berdasarkan teks pendek karena tidak memiliki konteks. Misalnya, "Film itu sama bagusnya dengan film terakhirnya" sepenuhnya bergantung pada apa pendapat orang tersebut tentang pemikiran model sebelumnya.

Keinginan pengguna terus berlanjut dan ketergantungan pada saran dan rekomendasi online yang diungkapkan oleh data hanyalah salah satu alasan di balik munculnya minat pada sistem baru yang berhubungan langsung dengan pendapat sebagai objek utama. Analisis sentimen berkonsentrasi pada sikap, sedangkan penambahan teks tradisional berfokus pada analisis fakta. Ada beberapa bidang penelitian utama yang mendominasi dalam analisis Sentimen: klasifikasi sentimen, klasifikasi sentimen berbasis Sentimen dan rangkuman opini. Klasifikasi sentiment dengan mengklasifikasikan seluruh dokumen sesuai opini pada objek tertentu. Klasifikasi sentimen berbasis fitur di sisi lain menganggap opini pada fitur objek tertentu. Kesimpulan opini berbeda dengan kesimpulan teks tradisional karena hanya fitur produk yang ditambang, dimana pelanggan telah mengungkapkan opini mereka.

Agar memudahkan, karya ilmiah ini disusun menjadi 6 bagian dalam pembahasan: Bagian 1 menyajikan sumber data yang digunakan untuk opini mining. Bagian 2 memperkenalkan pembelajaran mesin untuk klasifikasi sentimen. Bagian 3 menyajikan beberapa aplikasi klasifikasi sentiment dan tool apa saja yang digunakan untuk klasifikasi sentimen. Bagian 5 adalah tentang evaluasi kinerja yang dilakukan. Bagian terakhir menyimpulkan studi kami dan membahas beberapa saran untuk penelitian selanjutnya.

Analisis sentiment merupakan penelitian yang sedang tren di bidang Natural Language Processing (NLP). Hal ini dikarenakan kemudahan untuk memperoleh data. Analisis sentiment sering menggunakan informasi dari postingan – postingan atau berita yang ada di website media sosial. Salah satu contoh adalah website berita, twitter, dan facebook. Dengan menggunakan API yang sesuai dengan sosial media yang ditelitinya maka data dapat diambil selama periode tertentu. Analisis sentimen umumnya dilakukan pada teks yang berbahasa inggris

karena *tools*, *script*, kamus, dan *corpus* (kumpulan teks) yang sudah tersedia.

Bahasa Indonesia termasuk sebagai low resource language atau bahasa yang sumber dayanya rendah dikarenakan kurangnya *corpus*, kamus, *tools*, dan *script* yang mendukung untuk melakukan Pengolahan Bahasa Natural (NLP) [2]. Selain itu Bahasa Indonesia yang formal jarang digunakan di sosial media, blog, website, dan *chatting platform* (WA, Line, Kakao Talk) sehingga untuk mengolah teks yang berbahasa Indonesia menjadi tantangan bagi para peneliti yang tertarik untuk mengembangkan NLP yang berbahasa Indonesia yang lengkap, efektif, dan efisien.

Sentimen analisis yang menggunakan data berbahasa Indonesia masih sangat sedikit jika dibandingkan dengan sentimen analisis yang menggunakan Bahasa Inggris. Hal ini dikarenakan tools untuk mengolah teks berbahasa Indonesia seperti Stemmer, Pos-Tagger, dan Pengolah Kata Tidak Baku menjadi Baku masih sangat sedikit. Selain itu *corpus* yang berbahasa Indonesia dengan data yang sudah bersih dan siap untuk dilakukan prapengolahan, klasifikasi, atau penklasteran masih sangat jarang.

Kecenderungan peneliti NLP yang menggunakan Bahasa Indonesia untuk tidak membagikan dataset yang digunakan juga mempengaruhi jumlah hasil penelitian NLP yang menggunakan Bahasa Indonesia. Berdasarkan hasil studi pustaka diketahui bahwa Universitas Indonesia (UI) merupakan Perguruan Tinggi yang cukup serius untuk mengembangkan teknologi NLP yang menggunakan bahasa Indonesia. UI memiliki laboratorium yang bernama Information Retrieval Lab [3]. Beberapa tool yang sudah dikembangkan oleh Information Retrieval Lab adalah Indonesian Wordnet, Indonesian Dictionary, Indonesian TreeBank, dan PosTag Corpus. Perguruan tinggi lain juga berkontribusi terhadap penelitian NLP namun masih belum terorganisasi seperti pada Information Retrieval Lab. Selain Information Retrieval Lab terdapat beberapa institusi atau organisasi yang juga mengembangkan NLP khususnya sentiment analisis seperti Indonesia Indicator dan Pulse Labs Jakarta. Indonesia Indicator [4] merupakan perusahaan yang menyediakan jasa untuk mengolah data media menjadi pengetahuan dan prediksi sehingga dapat membantu dalam mengambil keputusan. Sedang Pulse Labs Jakarta [5] merupakan cabang dari Pulse Labs International yang bekerjasama dengan Bappenas (Badan Perencanaan Pembangunan Nasional) yang meneliti sosial media dan sumber media digital lain sehingga memperoleh pengetahuan atau informasi tentang perkembangan sosial di Indonesia.

Kedua hal tersebut yang menjadi masalah yg ingin diselesaikan pada penelitian ini. fokus membahas analisis sentiment pada bahasa inggris dan Indonesia. Untuk menyelesaikan masalah tersebut maka ditentukan tujuan penelitian, yaitu melakukan kajian pustaka secara tradisional yang mana sumber – sumber data diambil secara acak sesuai dengan konteks yang dikaji. Hasil

studi penelitian ini diharapkan dapat memberikan gambaran secara menyeluruh tentang analisis sentiment pada media online dan membantu para peneliti yang akan datang dalam menentukan topik penelitiannya.

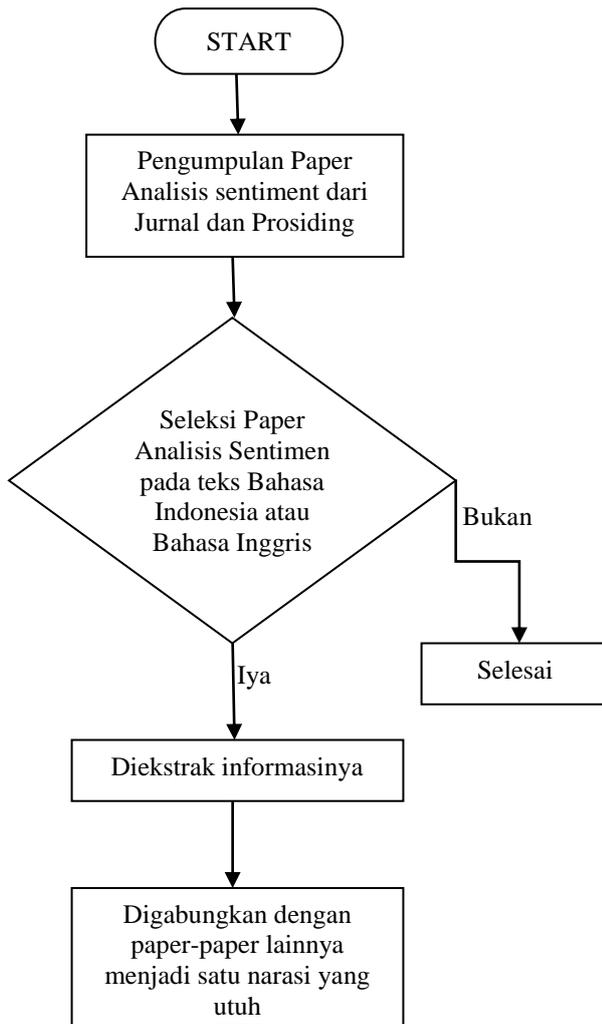
2. Tinjauan Pustaka

Survei dan review tentang analisis sentimen telah dilakukan berbagai peneliti seperti pada penelitian yang dilakukan Vinodhini dan Chandrasekaran [6] yang menjelaskan tentang analisis sentimen berbahasa Inggris dengan sub pokok bahasan meliputi asal data, metode klasifikasi, aplikasi dan tool, evaluasi performa, dan arah penelitian di masa depan. Tantangan utama yang diperoleh dari kajian pustaka yang dilakukan oleh Vinodhini dan Chandrasekaran adalah melakukan analisis sentimen untuk Bahasa lain, menyelesaikan masalah dengan ekspresi negasi, menghasilkan ringkasan dari opini berdasarkan atribut data, kompleksitas dari kalimat atau dokumen, dan atribut data yang implisit. Sedangkan pada penelitian yang dilakukan oleh Nanli et al [7] adalah mempelajari paper – paper tentang analisis sentiment berbahasa Cina dan menuliskannya menjadi sebuah kajian paper yang membahas tentang framework, ekstraksi ciri, aplikasi bisnis, dan sentiment analisis berbasis blog. Berdasar dari penelitian Nanli et al dikatakan bahwa penelitian sentiment analisis perlu didukung dengan data bahasa yang empiris. Selain itu disebutkan juga tidak ada analisis mengenai nilai komersial dari analisis sentimen untuk ulasan produk. Pada penelitian yang dilakukan oleh Gupta et al [8] terdapat ringkasan dari beberapa penelitian tentang analisis sentiment seperti analisis sentiment dengan pembelajaran mesin, analisis sentiment pada ulasan film, analisis sentiment dengan kata negasi, *dual sentiment analysis*, dan meringkas teks secara otomatis. Kemudian disimpulkan bahwa untuk memperoleh hasil analisis sentiment yang lebih baik (akurat dan efisien) perlu digunakan metode pembelajaran mesin yang lebih modern. Kemudian pada kajian pustaka yang dilakukan oleh Bansal dan Singh [9] dibuat tabel dengan kriteria sebagai berikut *approaches*, *type of data used*, *algorithms used*, *polarity*, dan *granularity of analysis*. Dari tabel tersebut diketahui bahwa algoritma yang paling sering digunakan adalah SVM dan Naïve Bayes. Sedangkan data yang paling sering digunakan adalah data IMDB dan Restoran Review. Kemudian dijelaskan juga tentang pengaplikasian dari analisis sentiment yaitu *online advertising*, *prediction system*, dan *competitive analysis*. Selanjutnya pada penelitian yang dilakukan oleh Shivaprasad dan Shetty [10] melakukan kajian tentang analisis sentimen pada komentar dari suatu produk. Pada penelitian ini dijelaskan tentang tingkatan analisis sentimen seperti analisis sentimen pada tingkat dokumen, tingkat kalimat, dan tingkat kata. Selain itu dijelaskan juga tentang proses analisis sentimen yang mana terdapat persiapan data, analisis review, dan klasifikasi sentimen. Terdapat pembahasan tentang metode dan teknik untuk menemukan polarity dari sentimen. Lalu pada penelitian Kaur [1] dijelaskan

tentang aplikasi dari pengertian, metode analisis sentiment, dan penelitian yang masih belum dilakukan. Dari penelitian Kaur diketahui bahwa ada banyak algoritma atau metode yang dapat digunakan untuk mengembangkan analisis sentimen. Dari hasil kajian teori yang dilakukan Kaur diketahui bahwa SVM dan Naïve Bayes merupakan algoritma yang paling populer. Mengekstrak konteks dari tulisan menjadi tantangan analisis sentimen karena perlu analisis yang lebih dalam untuk memperoleh konteks tulisan tersebut. Selain itu dataset seperti IMDB, flipkart, dan Amazon digunakan secara luas oleh para peneliti. Seperti halnya penelitian dengan menggunakan data twitter sangat populer di kalangan peneliti saat ini. Penelitian sebelumnya terdiri dari lima ulasan paper [6][8][9][10][1] yang membahas tentang analisis sentimen dengan Bahasa Inggris, sedangkan ulasan paper [11][12][13][14][15][16] membahas tentang analisis sentimen dengan Bahasa Indonesia. Selain itu untuk memberikan gambaran secara umum perkembangan analisis sentimen dengan Bahasa Inggris dan Bahasa Indonesia maka perlu dibuat sebuah karya ilmiah yang berisikan penjelasan secara menyeluruh mengenai posisi penelitian analisis sentimen dengan Bahasa Inggris dan Bahasa Indonesia, metode analisis sentimen, tren penelitian, dan aplikasi sentiment analisis di sektor riil.

3. Metode Penelitian

Perkembangan penelitian analisis sentimen dengan Bahasa Inggris sangat diminati oleh peneliti, karena *tools*, *script*, kamus, dan *corpus* (kumpulan teks) yang sudah tersedia. Sedangkan penelitian analisis sentimen dengan Bahasa Indonesia masih belum menjadi perhatian bagi para peneliti.. Oleh karena itu untuk merangkum paper – paper analisis sentimen yang tersebar di berbagai database Jurnal dan Seminar maka dilakukan penelitian yang mempelajari perkembangan, pertanyaan penelitian, tantangan, tren, dan riset di masa depan dari analisis sentimen dengan Bahasa Inggris dan Bahasa Indonesia yang kemudian dikumpulkan menjadi sebuah tulisan ilmiah yang mampu memberikan informasi secara menyeluruh tentang analisis sentimen dengan Bahasa Inggris dan Bahasa Indonesia seperti ditunjukkan pada Gambar 1. Paper – paper tersebut dikumpulkan dari berbagai sumber seperti Digital Library IEEEExplorer, Jurnal Nasional, prosiding Nasional, dan Seminar Nasional. Isi dari paper - paper rujukan dirangkum menjadi bagian dari sub pokok bahasan dari kajian paper ini. Sub pokok bahasan dari paper ini terdiri dari sumber data, metode, prapengolahan, tren penelitian, dan aplikasi riil.



Gambar 1. Alur Penelitian

4. Pembahasan

4.1 Sumber Data

Pendapat pengguna merupakan kriteria utama untuk peningkatan kualitas layanan yang diberikan dan peningkatan hasil. Blog, review situs dan mikroblog memberikan informasi yang banyak tentang tingkat penerimaan produk dan layanan.

1. Blog

Dengan semakin meningkatnya penggunaan internet, blog dan halaman blog semakin berkembang pesat. Halaman blog telah menjadi sarana yang paling populer untuk mengekspresikan pendapat pribadi seseorang. Blogger merekam kejadian sehari-hari dalam kehidupan mereka dan mengungkapkan pendapat, perasaan, dan emosi mereka di sebuah blog. Banyak blog ini berisi ulasan tentang banyak produk, isu, dll. Blog digunakan sebagai sumber opini dalam banyak penelitian yang berkaitan dengan analisis sentimen.

2. Tinjau situs

Bagi pengguna manapun dalam membuat keputusan pembelian, pendapat orang lain bisa menjadi faktor penting. Ringkasan ulasan pengguna yang besar dan berkembang tersedia di Internet. Ulasan untuk produk atau layanan biasanya didasarkan pada pendapat yang diungkapkan dalam format yang tidak terstruktur. Data resensi yang digunakan di sebagian besar kajian klasifikasi sentimen dikumpulkan dari situs web e-commerce seperti www.amazon.com (ulasan produk), www.yelp.com (ulasan restoran), www.CNET.com (produk review) dan www.reviewcentre.com, yang menampung jutaan ulasan produk oleh konsumen.

3. Micro-blogging

Twitter adalah layanan microblogging yang populer dimana pengguna membuat pesan status yang disebut "tweets". Tweet ini terkadang mengungkapkan pendapat tentang topik yang berbeda. Pesan tweet juga digunakan sebagai sumber data untuk mengklasifikasikan sentimen.

4.2 Klasifikasi Sentimen

Beberapa metode atau teknik yang digunakan untuk memperoleh opini dari dokumen yang dianalisis yaitu Naïve Bayes [12], Support Vector Machine (SVM) [13], Decision Tree [14], KNN [15], Random Forest [14], dan Rabin Karp [16].

4.3 Pra Pengolahan

Jenis prapengolahan yang digunakan dalam beberapa penelitian, dapat dirangkum ada 2 macam, yaitu *Preprocessing*, dilakukan untuk menghasilkan data bersih. *Preprocessing* meliputi: casefolding, filtering, tokenasi, slang replacement dan stopword removal. *Ekstraksi Fitur*, meliputi: Unigram, Negation, TF dan TF-IDF

4.4 Tool dan Library

Untuk mendapatkan sentiment analisis perlu menggunakan tool atau library. Sehingga memudahkan peneliti untuk fokus pada kontribusi yang ingin dicapainya. Pada penelitian sebelumnya digunakan tool WEKA, Rapid Miner dll.

4.5 Aplikasi Riil

Sentimen analisis menjadi salah satu analisis populer bagi industri data science di Indonesia. Hal ini diakibatkan permintaan pengguna seperti calon wakil rakyat, perusahaan, e-commerce, pemerintah, dan lain – lain. Contohnya beberapa kandidat wakil rakyat menggunakan jasa dari indonesi indicator untuk mengetahui elektabilitas di sosial media. Riset pemerintah seperti yang dilakukan oleh Pulse Lab Jakarta yang merupakan kerjasama antara United Nations and Badan Perencanaan Pembangunan Nasional (Bappenas). Pulse Lab melakukan riset tentang pengaruh kenaikan BBM dengan kenaikan harga bahan makanan.

5. Kesimpulan dan Saran

Deteksi sentimen memiliki beragam aplikasi dalam sistem informasi, termasuk mengklasifikasikan ulasan, meringkas review dan aplikasi real time lainnya. Ada kemungkinan banyak aplikasi lain yang tidak dibahas. Ditemukan bahwa klasifikasi sentimen sangat bergantung pada domain atau topik. Dari hasil di atas, jelas bahwa baik model klasifikasi secara konsisten tidak bekerja dengan baik, jenis fitur yang berbeda memiliki distribusi yang berbeda. Juga ditemukan bahwa berbagai jenis fitur dan algoritma klasifikasi dikombinasikan dengan cara yang efisien untuk mengatasi kekurangan masing-masing dan mendapatkan keuntungan dari satu sama lain, dan akhirnya meningkatkan kinerja klasifikasi sentimen.

Dipenelitian akan datang, diperlukan lebih banyak pekerjaan untuk memperbaiki performa algoritma untuk klasifikasi. Analisis sentimen bisa diterapkan untuk aplikasi baru. Meskipun teknik dan algoritma yang digunakan untuk analisis sentimen semakin cepat, namun banyak masalah dalam bidang ini yang belum terpecahkan. Aspek utama ada dalam penggunaan bahasa, yang berhubungan dengan ekspresi negasi; ekstraksi fitur/atribut produk, kompleksitas kalimat/dokumen, penanganan fitur produk implisit, dan lain-lain. Penelitian berikutnya semoga dapat menjawab tantangan ini..

Daftar Pustaka

- [1] H. Kaur, V. Mangat, and Nidhi, "A Survey of Sentiment Analysis techniques," in *International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017)*, 2017, pp. 921–925.
- [2] T. A. Le, D. Moeljadi, M. Yasuhide, and T. Ohkuma, "Sentiment Analysis for Low Resource Languages : A Study on Informal Indonesian Tweets," *12th Work. Asian Lang. Resour.*, pp. 123–131, 2016.
- [3] I. R. Lab, "Information Retrieval," *Information Retrieval Lab - Faculty of Computer Science, Universitas Indonesia*, 2016. .
- [4] I. Indicator, "Indonesia Indicator," 2017. .
- [5] P. L. Jakarta, "Pulse Lab Jakarta," 2016. .
- [6] G. Vinodhini and R. Chandrasekaran, "Sentiment Analysis and Opinion Mining: A Survey," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 2, no. 6, pp. 282–292, 2012.
- [7] Z. Nanli, Z. Ping, L. Weiguo, and C. Meng, "Sentiment analysis: A literature review," *2012 Int. Symp. Manag. Technol.*, pp. 572–576, 2012.
- [8] P. Gupta, R. Tiwari, and N. Robert, "Sentiment Analysis and Text Summarization of Online Reviews: A Survey," in *International Conference on Communication and Signal Processing*, 2016, pp. 241–245.
- [9] N. Bansal and A. Singh, "A review on opinionated sentiment analysis based upon machine learning approach," *2016 Int. Conf. Inven. Comput. Technol.*, pp. 1–6, 2016.
- [10] S. T. K and J. Shetty, "Sentiment Analysis of Product Reviews: A Review," in *International Conference on Inventive Communication and Computational Technologies*, 2017, pp. 298–303.
- [11] P. Aliandu, "Sentiment Analysis to determine Accommodation , Shopping and Culinary Location on Foursquare in Kupang City," *Procedia - Procedia Comput. Sci.*, vol. 72, pp. 300–305, 2015.
- [12] Franky and R. Manurung, "Machine Learning-based Sentiment Analysis of Automatic Indonesian Translations of English Movie Reviews," *Int. Conf. Adv. Comput. Intell. Its Appl.*, vol. 1, 2008.
- [13] H. Wijaya, A. Erwin, A. Soetomo, and M. Galinium, "Twitter Sentiment Analysis and Insight for Indonesian Mobile Operators," *Inf. Syst. Int. Conf.*, no. December, p. 367, 2013.
- [14] Zamahsyari and A. Nurwidyantoro, "Sentiment analysis of economic news in Bahasa Indonesia using majority vote classifier," *Proc. 2016 Int. Conf. Data Softw. Eng. ICoDSE 2016*, 2017.
- [15] S. K. Lidya, O. S. Sitompul, and S. Efendi, "Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine (Svm)," *Semin. Nas. Teknol. dan Komun. 2015*, vol. 2015, no. Sentika, pp. 1–8, 2015.
- [16] Y. Astuti and S. Adi, "Information Retrieval Mendeteksi Konten Anarkis Pada Web Keagamaan Menggunakan Algoritma Rabin Karp," vol. 17, no. 4.

Biodata Penulis

Sumarni Adi, memperoleh gelar Sarjana Komputer (S.Kom), pada Jurusan Teknik Informatika STMIK Amikom Yogyakarta, pada tahun 2009. Memperoleh gelar *Master Of Computer Science (M.Cs)* pada program Ilmu Komputer Universitas Gadjah Mada Yogyakarta pada tahun 2014. Penelitian yang ditekuni dibidang Data Mining dan NLP. Saat ini sebagai Staf Pengajar di Program Studi Informatika Universitas Amikom Yogyakarta.

Murni Wulandari, saat ini merupakan mahasiswa semester 5 pada Fakultas Ilmu Komputer, Program Studi Informatika di Universitas Amikom Yogyakarta

Anis Kemala Mardiana, saat ini merupakan mahasiswa semester 5 pada Fakultas Ilmu Komputer, Program Studi Informatika di Universitas Amikom Yogyakarta

Ahmad Muzakki, saat ini merupakan mahasiswa semester 5 pada Fakultas Ilmu Komputer, Program Studi Informatika di Universitas Amikom Yogyakarta

