

TEKNOLOGI *OPINION MINING* UNTUK Mendukung *STRATEGIC PLANNING*

Dwi Rolliawati*¹, Khalid², Indri Sudanawati Rozas³

^{1,2,3}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, UIN Sunan Ampel Surabaya
Email: ¹dwi_roll@uinsby.ac.id, ²khalid@uinsby.ac.id, ³indrisrozas@uinsby.ac.id

*Penulis Korespondensi

(Naskah masuk: 23 Januari 2019, diterima untuk diterbitkan: 06 Februari 2020)

Abstrak

Banjir data di era Big Data sudah tidak bisa terelakkan lagi. Termasuk di dalamnya data yang sangat melimpah di media sosial daring. Peluang inilah yang ditangkap sebagai alasan utama pada penelitian ini. *Opinion mining* sebagai salah satu teknologi dalam mengolah data teks untuk memperoleh arah informasi dari komentar/opini masyarakat. Mengambil obyek penelitian UIN Sunan Ampel Surabaya, penelitian ini bertujuan untuk menganalisis opini masyarakat tentang kampus Islam terbesar di Surabaya. Sehingga bisa menjadi pendukung keputusan bagi pihak manajemen untuk merumuskan perencanaan strategis terwujudnya visi *World Class University*. Penelitian ini menggunakan 4009 data sampel berbahasa Indonesia yang diambil dari opini masyarakat di media sosial Twitter dalam kurun waktu dua tahun terakhir (2017 – 2018). Dari 4009 data dihasilkan 31837 jenis kata setelah melalui proses *stop-word removal*. Berdasarkan analisis *sentiment* menggunakan pendekatan Vader dan Liu yang divisualisasikan melalui grafik K-Means, dihasilkan bahwa opini publik terhadap UIN Sunan Ampel mengarah pada sentimen 'netral' sebesar 97,54%, sedangkan *sentiment positif* = 2,16%, dan *sentiment negatif* = 0,34%. Hasil tersebut membuktikan bahwa *Information Capital* tentang UIN Sunan Ampel perlu diperkuat menuju nilai "positif". Sehingga diperlukan upaya maksimal untuk membangun *innovation and commercially supremacy, perception (public relation)* dan *scalability strategies* supaya *internal operation* bisa handal untuk ketercapaian visi misi UIN Sunan Ampel Surabaya.

Kata Kunci: *Big Data, Opinion Mining, World Class University, Twitter, K-Means, Vader, Liu.*

OPINION MINING TECHNOLOGY TO SUPPORT STRATEGIC PLANNING

Abstract

Data deluge in Big Data era is inevitable, this including a very abundant data in online social media. This phenomenon was chosen as the main background reason in this research. Opinion mining is as one of the technologies in processing text data to obtain information direction from public comments/opinions. Taking the object of research at Sunan Ampel Islamic State University Surabaya, this study aims to analyze public community opinion toward the biggest Islamic campus in Surabaya. Hopefully, it would be beneficial as decisional support for management in formulating strategic planning to manifest the World Class University vision. This study uses 4009 Indonesian language sample data taken from public opinion on Twitter social media in the past two years (2017 - 2018). Out from 4009 data, 31837 types of words are obtained after going through a stop-word removal process. Based on sentiment analysis by Vader and Liu's approach which was visualized by K-Means graphs, the finding was that 97,54% of public opinion toward Sunan Ampel Islamic State University Surabaya led to a 'neutral' sentiment, while positive = 2,16% and negative = 0,34%. These results prove that Information Capital about Sunan Ampel UIN needs to be strengthened towards "positive" image. For this reason, maximum effort is needed to build innovation and commercialization of supremacy, perception (public relations) and scalability strategies so that internal operations can be reliable in achieving the vision of Sunan Ampel Islamic State University Surabaya.

Keywords: *Big Data, Opinion Mining, World Class University, Twitter, K-Means, Support Vector Machine.*

1. PENDAHULUAN

Teknologi media sosial telah menjadi mainstream bagi peradaban manusia, dan bukan lagi sesuatu yang baru bagi masyarakat dunia, termasuk di Indonesia. Berdasarkan penelitian yang dilakukan oleh *We Are Social* yang bekerjasama dengan Hootsuite, diungkapkan bahwa sebesar 130 juta *Are Social* total populasi Indonesia saat ini mencapai 265,4 juta jiwa, dan setengah dari populasi tersebut (132,7 juta jiwa) merupakan pengguna internet (Hutchison, 2018).

Fakta lainnya menunjukkan bahwa orang Indonesia rata-rata menghabiskan waktu selama 8,85 jam atau setara dengan 8 jam 51 menit untuk berselancar di internet. Adapun rata-rata berkecimpung di media sosial melalui berbagai perangkat hingga 3,38 jam (3 jam 23 menit). Dari berbagai macam media sosial yang ada saat ini, *platform* yang paling banyak digunakan oleh orang Indonesia, di antaranya *YouTube* sebesar 43%, *Facebook* sebesar 41%, *WhatsApp* sebesar 40%, *Instagram* sebesar 38%, *Line* sebesar 33%, *BBM* sebesar 28%, *Twitter* sebesar 27%, *Google+* sebesar 25%, *FB Messenger* sebesar 24%, *LinkedIn* sebesar 16%, *Skype* sebesar 15%, dan *WeChat* sebesar 14% (Hutchison, 2018).

Hal tersebut menunjukkan sebagian orang lebih banyak bersosialisasi di dunia maya daripada dunia nyata. Fungsi awal dari media sosial, yang tadinya hanya untuk mendekatkan dan menjalin silaturahmi antar teman atau keluarga, kini telah berkembang menjadi penyedia dan penyebar informasi, segmen pasar baru bagi dunia bisnis, serta pengembangan jejaring, komunitas, penggirirangan opini dan riset.

Teknologi *opinion mining* atau *sentiment analysis* sebagai salah satu dari *natural language processing* atau studi komputasional yang mengumpulkan, menganalisis, menelusuri opini, sikap dan emosi publik terhadap suatu produk atau entitas tertentu (Medhat, Hassan and Korashy, 2014). Melalui sumber daya media sosial telah banyak dikembangkan penelitian terkait *opinion mining*, salah satunya adalah Lie Chen, dkk yang memetakan tren riset dan sitasinya tentang *market stock* dari web media. Hasil riset Lie menunjukkan bahwa sitasi paling tinggi pada riset *Twitter mood predicts the stock market* dengan besaran 2785 sitasi. Hasil penelitian ini menunjukkan bahwa web media terutama social media terutama Twitter bisa digunakan untuk memprediksi tren suatu entitas (Li *et al.*, 2018).

Berbeda dengan Lie, El-Gohary menggali opini *stakeholder* sebagai alat identifikasi dalam membuat *large-scale highway project* (Lv and El-Gohary, 2016). Sedangkan Thamrin membuat aplikasi matriks *rule based SWOT analysis* menggunakan teknik *sentiment analysis*. Metode sentiment analysis digunakan untuk menjangkir nilai sentiment dari opini stakeholder suatu perguruan tinggi berdasar pada Standar Nasional Pendidikan Tinggi (Thamrin and Pamungkas, 2017). Melalui teknologi *opinion mining* tersebut akan diketahui *positioning* suatu perguruan tinggi menurut persepsi masyarakat. Dengan demikian urgensi dari penelitian ini masuk kategori penting mengingat sumber daya media sosial di Indonesia sangatlah besar dan belum banyak riset yang mengangkat *opinion mining* sebagai langkah untuk menentukan *strategic planning* bagi perguruan tinggi. Adapun yang menjadi obyek penelitian adalah UIN Sunan Ampel Surabaya. Hal ini dimaksudkan untuk mengetahui apakah tercapainya visi misi UIN Sunan Ampel Surabaya “menjadi Universitas Islam yang unggul dan kompetitif bertaraf internasional” sudah bisa diprediksi secara tepat dan tidak hanya sebagai wacana saja.

1. STATE OF THE ART

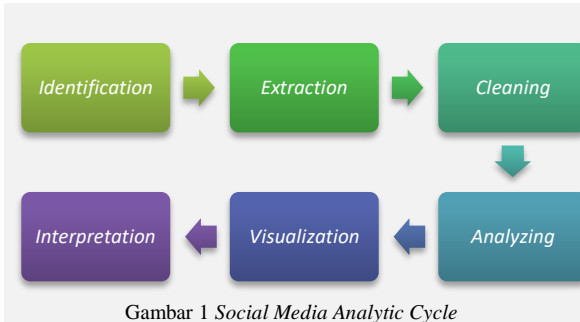
A. Literature Review

Opinion Mining

Pada beberapa literatur (Pang & Lee, 2006), (Liu, 2012), dan (Sneka & Vidhya, 2016), disebutkan bahwa analisis sentimen sama dengan *opinion mining*, yaitu bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan berbagai atribut lainnya. Berdasarkan literatur ditemukan penyebutan berbagai nama dengan perbedaan definisi yang sangat tipis seperti *sentiment analysis*, *opinion mining*, *opinion extraction*, *sentiment mining*, *subjectivity analysis*, *affect analysis*, *emotion analysis*, dan *review mining*. Semua istilah tersebut berada di payung besar *sentiment analysis* dan *opinion mining*. Makna opini itu sendiri masih sangat luas. Analisis sentimen dan penambangan opini berfokus pada pendapat yang mengungkapkan atau menyiratkan sentimen positif atau negatif. Analisis data pada sentiment menghasilkan *tuple* berpasangan (Polaritas, Subjektivitas). *Tuple* Polaritas memiliki nilai [-1, 1] ==> negative~positive. Sentimen *Subjectivity* memiliki nilai antara 0 sampai 1, dimana 0 paling objektif dan 1 paling subjektif (Liu, 2012).

Social Media Analytic (SMA)

SMA adalah sebuah proses pengumpulan data dari media sosial dan analisisnya untuk mendapatkan '*insights*' atau informasi berharga untuk suatu tujuan tertentu (Stieglitz *et al.*, 2018). Terdapat 6 langkah dalam *social media analytic* (Dr. Khan, 2017) yaitu *identification*, *extraction*, *cleaning*, *analyzing*, *visualization* dan *interpretation* yang digambarkan sebagaimana pada Gambar 1.



Gambar 1 Social Media Analytic Cycle

Tantangan dalam SMA adalah (Stieglitz *et al.*, 2014):

1. Pendek (*short in lengths*): bahkan terkadang tidak mengandung sebuah kalimat yang utuh menurut tata bahasa (*grammar*).
2. *Noise*: data media sosial penuh dengan *noise* seperti *typos* (salah ketik), *encoding* yang tidak jamak, *slang*, dsb.
3. *Temporal*: informasi yang sedang *trending* biasanya hanya sesaat, sehingga SMA diharapkan dilakukan dengan cepat menggunakan model-model/teknik-teknik analisa data yang efisien.
4. *High-dimensional*: data di media sosial (teks, gambar, video, suara, dsb) adalah data tidak terstruktur berdimensi tinggi.
5. *Fine-grained*: data di media sosial berasal dari banyak user yang masing-masingnya bisa jadi membahas beberapa topik yang berbeda. Sehingga komunitas (kelompok), topik, maupun klasifikasi yang ada menjadi besar (*fine-grained*).
6. *Large in volume & high velocity*: data yang sangat besar dan bertambah besar dengan cepat.
7. *A lot of external information*: informasi terkadang lebih banyak terkandung dari luar (eksternal) seperti *url website*, video, atau hal lain yang dibagikan oleh pengguna media sosial.

Pendekatan Vader dan Liu

Metode VADER merupakan singkatan dari *Valence Aware Dictionary for sEntiment Reasoning* yang menggunakan *simple rule based model*. Dari riset yang dilakukan, metode VADER menunjukkan tingkat akurasi yang tinggi untuk sentiment menggunakan data tweet, yaitu sebesar 0,96 jika dibandingkan dengan 11 metode lainnya sebagaimana Gambar 2.

Sedangkan metode Liu yang digagas oleh Bing Liu menggunakan pendekatan *supervised pattern learning* untuk mengekstraksi fitur objek dan mengidentifikasi orientasi suatu opini. Untuk mengidentifikasi orientasi opini, Liu menggunakan pendekatan berbasis leksikon. Pendekatan leksikon pada dasarnya menggunakan kata-kata dan frasa dalam sebuah kalimat untuk menentukan sentiment dari sebuah opini. Pendekatan berbasis leksikon memiliki langkah: identifikasi kata-kata opini, mencari kata-kata negasi, dan terakhir but-clauses (Mishra, 2012).

| | 3-Class Classification Accuracy (F1 scores) | | | |
|----------------|---|-------------|-------------|-------------|
| | Test Sets | | | |
| | Tweets | Movie | Amazon | NYT |
| VADER | 0.96 | 0.61 | 0.63 | 0.55 |
| NB (tweets) | 0.84 | 0.53 | 0.53 | 0.42 |
| ME (tweets) | 0.83 | 0.56 | 0.58 | 0.45 |
| SVM-C (tweets) | 0.83 | 0.56 | 0.55 | 0.46 |
| SVM-R (tweets) | 0.65 | 0.49 | 0.51 | 0.46 |
| NB (movie) | 0.56 | 0.75 | 0.49 | 0.44 |
| ME (movie) | 0.56 | 0.75 | 0.51 | 0.45 |
| NB (amazon) | 0.69 | 0.55 | 0.61 | 0.48 |
| ME (amazon) | 0.67 | 0.55 | 0.60 | 0.43 |
| SVM-C (amazon) | 0.64 | 0.55 | 0.58 | 0.42 |
| SVM-R (amazon) | 0.54 | 0.49 | 0.48 | 0.44 |
| NB (nyt) | 0.59 | 0.56 | 0.51 | 0.49 |
| ME (nyt) | 0.58 | 0.55 | 0.51 | 0.50 |

Gambar 2 Perbandingan Akurasi Metode VADER (Hutto and Gilbert, 2014)

Strategic Planning

Perencanaan strategis (*strategic planning*) adalah sebuah cara organisasi dalam mempersiapkan perubahan dan memberikan arahan serta mengkoordinasikan kegiatannya secara internal untuk mencapai tujuan organisasi yang ditetapkan (Ritson, 2011). Strategi sendiri dikategorikan ke dalam beberapa jenis. Ada *planned strategy*, *emergent strategy*, *opportunistic strategy*, dan *imposed strategy*. *Strategic plan* merupakan bentuk dari *planned strategy* yang ketika diturunkan ke dalam *deliberate strategy* yang diharapkan dapat diimplementasikan atau *realized strategy* (Ritson, 2011). Ritson menggambarkan hubungan antar jenis strategi dalam Gambar 3.

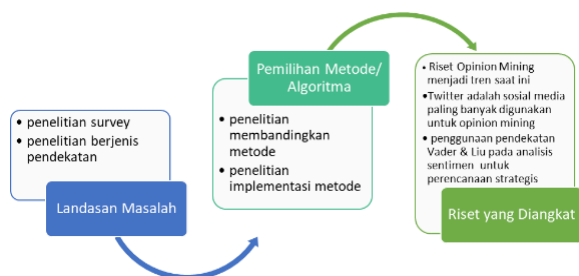


Gambar 3 Relasi Antar Jenis Strategi

B. Previous Research

Penelitian tentang *opinion mining* telah banyak dilakukan dan menjadi *trend* di era *big data* saat ini. "Apa yang orang lain pikirkan" selalu

menjadi bagian informasi penting bagi sebagian besar dari kita selama proses pengambilan keputusan (Pang and Lee, 2006). Relevansi penelitian terdahulu dengan penelitian ini digambarkan keterkaitannya dalam Gambar 4.



Gambar 4 Keterkaitan Penelitian Terdahulu

Beberapa penelitian yang menjadi tolok ukur landasan masalah dalam penelitian ini sebagaimana paparan berikut. Penelitian oleh Li, dkk yang memetakan 2785 tren riset dan sitasinya tentang stok market dan web media *Twitter* (Li *et al.*, 2018). Survei dari penelitian Hussein membahas faktor penting dan efek tantangan dalam analisis sentimen dari 47 makalah yang diteliti. (Singh and Husain, 2014).

Adapun terkait metode pengolahan data, Tahun 2014 Medhat melakukan survey tentang *opinion mining* dan *sentiment analysis* untuk menentukan tren riset dan algoritma yang digunakan. Dari hasil survei tersebut dihasilkan bahwa *Naïve Bayes* dan *Support Vector Machine* (SVM) adalah metode yang paling baik untuk diimplementasikan dalam *sentiment analysis* atau *opinion mining* dan *sentiment classification* (Medhat, Hassan and Korashy, 2014). Riset yang dilakukan McCoy, dkk untuk memeriksa dan memberi peringkat dari 264 universitas AS yang diekstrak dari keanggotaan dan daftar NCAA Divisi I dan diterbitkan dalam Berita AS menunjukkan bahwa *Times Higher Education* (THE), Peringkat Akademik Universitas Dunia (ARWU), dan *Money Magazine* menggunakan peringkat reputasi yang disesuaikan dan dibandingkan dengan skor keterlibatan Twitter di perguruan tinggi (McCoy, Nelson and Weigle, 2017).

Samir menggunakan pendekatan semantik untuk menemukan sikap pengguna dan wawasan bisnis dari *media social Twitter* dalam bahasa Arab, baik standar maupun dialek. Dalam penelitian tersebut dibuat pula versi pertama dari *Sentiment Ontology* (ASO) yang berisi kata-kata berbeda yang mengungkapkan perasaan dan seberapa kuat kata-kata mengungkapkan perasaan (Tartir and Abdul-Nabi, 2017). Janssen, dkk membuat studi kasus untuk membuktikan hipotesis bahwa sentimen *Twitter*

dapat secara potensial mendukung sistem peringkat universitas dengan menganalisis pernyataan yang diposting dan pendapat siswa dan guru dalam konteks lembaga pendidikan tinggi (Janssen *et al.*, 2015). Berdasarkan paparan riset sebelumnya terlihat nyata bahwa social media, salah satunya *Twitter* adalah sumber daya yang paling banyak digunakan untuk riset *opinion mining*.

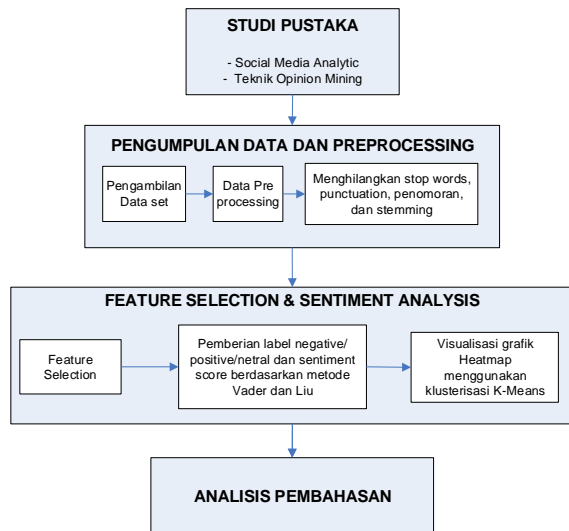
Adapun pijakan dalam pemilihan metode yang terbaik dalam penelitian ini adalah berdasarkan riset-riset berikut ini. Analisis *opinion mining* terhadap suatu produk di *amazon.com* dengan 3 metode classifier yaitu *naïve bayes*, SVM dan *decision tree*. Dari ketiga metode yang diujikan diperoleh tingkat akurasi yang tinggi sebesar 81,75 % terhadap penggunaan metode SVM dibandingkan dengan metode lainnya (Singla, Randhawa and Jain, 2017). Hal serupa dilakukan oleh (Devika, Sunitha and Ganesh, 2016) yang membandingkan metode-metode penggalian data dengan menggunakan algoritma *Support Vectore Machine*, N Gram, *Naïve Bayes*, KNN, ME Classifier, Feature Selection, dan Multilingual SA. Riset *opinion mining* (Science and Engineering, 2016) yang membandingkan algoritma *Support vector machine- nearest neighbour*, K-means *clustering* dan SOM (*Artificial Neural network*). Masih tentang *sentiment analisis/opinion mining*, dimana data opini dari *amazon.com* digunakan untuk kategorisasi suatu produk menggunakan *naïve bayes classifier*, *random forest* dan SVM (*support vector machine*) (Fang and Zhan, 2015).

Putranti mengimplementasi *sentiment public* dalam Bahasa Indonesia untuk mengetahui *positioning* telkomsel, indosat, iklan dan operator seluler lainnya (Putranti and Winarko, 2014) dan hal serupa dilakukan oleh Haddi (Haddi, Liu and Shi, 2013). Riset *opinion mining* oleh Rozi, dkk digunakan untuk menganalisis opini publik pada perguruan tinggi Polinema, Malang (Rozi, Pramono and Dahlan, 2012). Sisi lain *opinion mining* adalah untuk membandingkan produk dari hasil *review* konsumen untuk menentukan SWOT suatu produk dalam mendukung kesuksesan marketing (Auinger and Fischer, 2008).

Berdasarkan semua paparan riset sebelumnya, *state of the art* dalam penelitian ini adalah mengangkat *opinion mining* untuk mendukung perencanaan strategis suatu Lembaga Perguruan Tinggi dengan menggunakan pendekatan Vader dan Liu yang divisualisasikan dalam grafik *heatmap* berdasarkan metode K-Means.

2. METODOLOGI PENELITIAN

Secara umum, metodologi yang digunakan dalam menyelesaikan penelitian sebagaimana digambarkan pada Gambar 4.



Gambar 4. Metodologi Penelitian

Sebagaimana diagram alir pada Gambar 4, ada tiga langkah utama yang dilakukan dalam penelitian ini. Yang pertama yaitu studi pustaka, yang hasilnya telah dipaparkan pada bagian 2 (*State Of The Art*). Langkah kedua yaitu pengumpulan data dan *pre-processing*. Yang dilakukan pada langkah ini yaitu teknik pengumpulan data melalui *crawling* data Twitter secara daring dengan menggunakan API Key Twitter. Data yang diambil adalah 2 tahun terakhir yaitu 2017 dan 2018 dengan keyword yang ditentukan.

Langkah ketiga adalah *feature selection* dan *sentiment analysis*. Tahapan ini memetakan data tweet hasil *crawling* dengan memberikan label pada suatu kalimat *tweet* berdasarkan ukuran polaritas dan subyektifitasnya menggunakan pendekatan Vader & Liu. Untuk memudahkan dalam melakukan analisis dan menghilangkan data yang mirip/redundan maka dilakukan klasterisasi menggunakan metode K-Means. Hasil klasterisasi dokumen direpresentasikan dalam bentuk grafik *heatmap*. Setelah semua dilaksanakan barulah kemudian didapatkan hasil dan dilakukan pembahasan.

Langkah keempat yaitu analisis dan pembahasan, pada tahap ini dilakukan analisis hasil penelitian dalam mendukung perencanaan strategis. Perencanaan strategis yang dimaksud adalah perencanaan strategis IT.

3.1. Pengumpulan Data dan Pre-processing

Sebagaimana telah disebutkan sebelumnya, data yang di-*crawling* adalah 2 tahun terakhir (2017 dan 2018) dengan menggunakan *keyword* sebagai berikut:

1. “UINSA”,
2. “UINSBY”
3. “Sunan Ampel”
4. “UIN Sunan Ampel”
5. “UIN Sunan Ampel Surabaya”

Setelah di-*crawling*, diperoleh data mentah sebanyak 4009 *tweet*. Dan setelah dilakukan *pre-processing*, dari 4009 data *tweet* menghasilkan data *token* sebanyak 58937 kata dengan total tipe 10100 kata. Dari data 4009 *tweet* (58937 kata) menjadi 31837 kata ketika proses *stop-word removal*.

Untuk melihat detail frekuensi dari kata hasil *crawling* disimpan dalam bentuk tabulasi. Tabel 1 menampilkan frekuensi dari 20 kata teratas (dari 4009) yang muncul berdasarkan data opini publik tentang UIN Sunan Ampel Surabaya.

Tabel 1. Frekuensi Kata dari Data *Tweet*

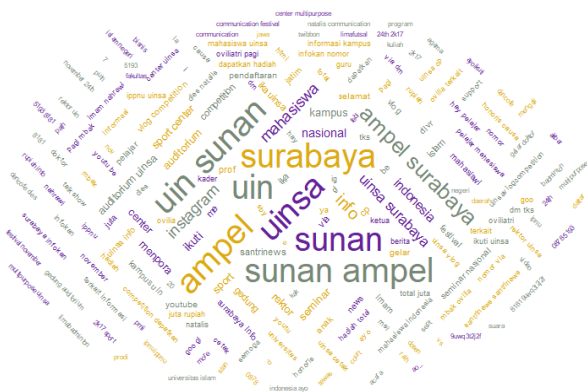
| No | Word | Word Count |
|----|----------------|------------|
| 1 | uinsa | 3325 |
| 2 | surabaya | 1380 |
| 3 | uin | 1317 |
| 4 | sunan | 1298 |
| 5 | ampel | 1289 |
| 6 | sunan ampel | 1286 |
| 7 | uin sunan | 1236 |
| 8 | ampel surabaya | 747 |
| 9 | info | 510 |
| 10 | instagram | 429 |
| 11 | mahasiswa | 429 |
| 12 | uinsa surabaya | 392 |
| 13 | indonesia | 315 |
| 14 | nasional | 293 |
| 15 | ikuti | 276 |
| 16 | auditorium | 257 |
| 17 | kampus | 237 |
| 18 | center | 232 |
| 19 | sport | 229 |
| 20 | rektor | 224 |

Jika Tabel 1 memberikan informasi rinci frekuensi kata, maka *word-cloud* memberikan model visual yang menggambarkan secara global apa saja kata yang paling banyak ditemukan. Kata yang frekuensinya lebih banyak atau paling sering muncul digambarkan dengan font yang lebih besar dibandingkan yang frekuensinya lebih sedikit. *Word-cloud* tabel di atas terdapat pada Gambar 5.

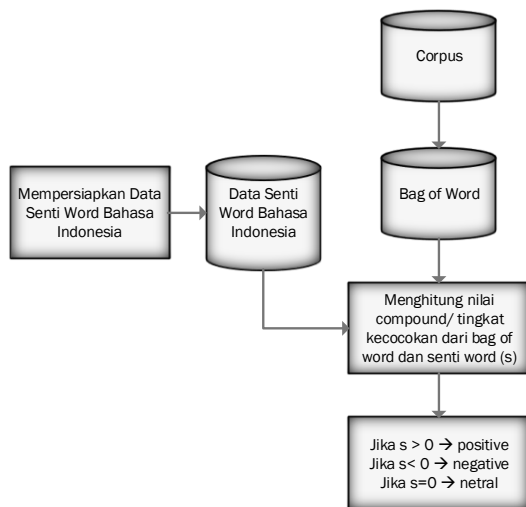
3.2. Feature Selection dan Sentiment Analysis

Tahapan ini memetakan data 4009 data *tweet* dengan memberikan label pada suatu kalimat

tweet berdasarkan ukuran polaritas dan subyektifitasnya menggunakan pendekatan dari Vader dan Liu ('Orange3 Text Mining Documentation', 2018). Vader memetakan opini publik dalam 3 kategori yaitu sentiment positif, sentiment negative dan netral. Nilai *compound* pada metode Vader ini adalah hasil campuran/gabungan atau bisa juga dikatakan sebagai rata-rata. Sedangkan metode Li hanya memetakan opini dalam 2 kluster yaitu positif dan *negative*. Secara detail, tahapan klusterisasi positif, negatif dan netral diilustrasikan pada Gambar 6.



Gambar 5. Word-cloud



Gambar 6. Tahapan Klusterisasi Opini

Berdasarkan tahapan metode Vader sebagaimana Gambar 6, dihasilkan nilai sentimen pada Tabel 2. Jika dilihat dari 50 data *corpus* (N=4009) yang ditampilkan pada Tabel 2, sebagian besar data *tweet* menghasilkan nilai sentimen absolut netral yaitu sama dengan “1”. Walaupun ada beberapa yang bernilai netral positif sebagaimana pada data *corpus* ke-15 yang memiliki nilai positif sebesar 0,266 dan netral = 0,734.

Hasil sentiment pada Tabel 2 menunjukkan bahwa sentiment positif = 2,16%, sentiment negatif = 0,34% dan sentiment netral = 97,54%. Hasil Tabel 2

masih ditemui beberapa data yang sama/redundansi sehingga 4009 data tweet yang sama di-*merger* tingkat kedekatan dokumennya menjadi 23 kluster dengan menggunakan algoritma K-Means. Hasil dari olahan K-Means divisualisasikan dalam bentuk grafik *HeatMap* seperti Gambar 7.

Tabel 2. Hasil Sentimen Analysis

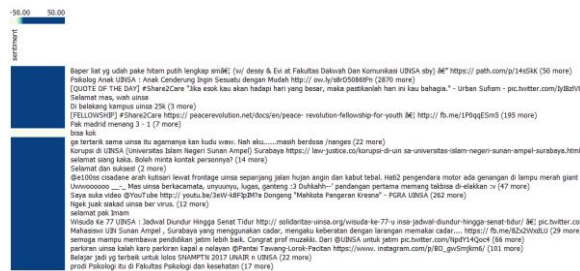
| No.id | Content | pos | neg | neu | compound |
|-------------|-------------------|--------|--------|--------|----------|
| 1 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 2 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 3 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 4 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 5 | http:// TRIBUNS | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 6 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 7 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 8 | #SURABAYA -l | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 9 | #BOJONEGOR | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 10 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 11 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 12 | #MAGETAN RA | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 13 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 14 | Psikolog Anak U | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 15 | Wonderful 2016 | 0,2660 | 0,0000 | 0,7340 | 0,6140 |
| 16 | Pementasan Seni | 0,1110 | 0,0000 | 0,8890 | 0,2500 |
| 17 | Pementasan Seni | 0,1110 | 0,0000 | 0,8890 | 0,2500 |
| 18 | [CALL FOR PAI | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 19 | #Repost @urban | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 20 | Last Meeting Wi | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 21 | Sambang Kader ' | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 22 | Saya suka video | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 23 | Pementasan Seni | 0,1110 | 0,0000 | 0,8890 | 0,2500 |
| 24 | 'yg gaada barisan | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 25 | [CONFERENCE | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 26 | [CONFERENCE | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 27 | [CONFERENCE | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 28 | Pementasan Seni | 0,1050 | 0,0000 | 0,8950 | 0,2500 |
| 29 | Pementasan Seni | 0,1110 | 0,0000 | 0,8890 | 0,2500 |
| 30 | Alhamdulillah se | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 31 | Owh kmu kuliah | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 32 | At UIN Sunan Ar | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 33 | Banyak cara untu | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 34 | UINSA di petang | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 35 | uinsa mbaak heh | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 36 | The English Ren | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 37 | HIMASI UNTAC | 0,1850 | 0,0790 | 0,7360 | 0,4329 |
| 38 | Pengurus Lubang | 0,8190 | 0,0000 | 0,7810 | 0,4215 |
| 39 | #nemu Foto 4taf | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 40 | Pertemuan +- 5 i | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 41 | Pertemuan +- 5 i | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 42 | #pathdaily (at UI | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 43 | Just posted a phc | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 44 | [GREETINGS] # | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 45 | [GREETINGS] # | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 46 | I liked a @YouTi | 0,2590 | 0,0000 | 0,7410 | 0,4215 |
| 47 | Joss Gandoss UI | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 48 | Anak Pedagang / | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 49 | Surabaya, Univer | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| 50 | ANAK PEDAG/ | 0,0000 | 0,0000 | 1,0000 | 0,0000 |
| AVERAGE | | 0,0216 | 0,0034 | 0,9754 | 0,0434 |
| AVERAGE (%) | | 2,16% | 0,34% | 97,54% | 4,34% |
| N= 4009 | | | | | |

Berdasarkan visualisasi tersebut, terlihat jelas bahwa semua data tweet yang diambil dari Twitter berdasarkan kata kunci UINSA, UINSBY, UIN Sunan Ampel, Sunan Ampel berisi opini/komentar yang bersifat netral. Gradasi warna dari setiap kolom nilai pos (positif), neg (negative) dan neu (netral) memiliki arti semakin berwarna gelap maka nilainya mendekati satu (1.0), semakin

berwarna muda maka nilainya semakin mendekati nol (0.0). Dan jika menggunakan metode Liu, yang hanya memiliki penilaian sentiment bernilai positif dan negative saja, grafik *HeatMap* akan terlihat sebagaimana Gambar 8.



Gambar 7. Grafik *HeatMap* dari hasil klusterisasi menggunakan *K-Means*(Lee and Park, 2012)



Gambar 8. Grafik *HeatMap* Metode Liu

Dengan nilai sentiment negatif = -50 dan nilai sentiment positif = 50, hal ini menunjukkan bahwa nilai sentiment mengarah pada netral dengan adanya bobot yang sama antara positif dan negatif.

3. HASIL DAN PEMBAHASAN

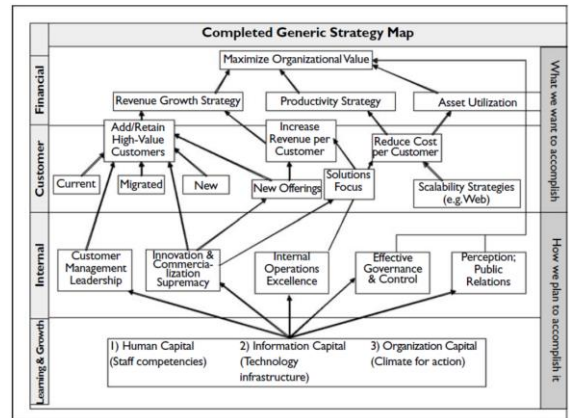
Strategic planning bukanlah sebuah proses yang sederhana, namun berupa fase kompleks dari manajemen strategis yang tergantung pada berbagai informasi terkait organisasi. Dan *balanced scorecard* memberikan konsep yang menyeluruh untuk kerangka berfikir menuju *strategic planning* (Stefanovska and Soklevski, 2014) *Balanced scorecard* memiliki empat perspektif: *finance, costumers, internal processes, learning and growth*.

Peta strategi (*Strategy map*) adalah alat sederhana yang mudah digunakan untuk menerjemahkan strategi dari visi dan misi organisasi ke dalam *strategic plan*. Dengan *balanced scorecard*, pembuatan *strategy map* akan lebih mudah dan menyeluruh (Hafnika et al., 2016).

Gambar 9 di bawah ini adalah model dasar dari *strategy map* berdasarkan *balanced scorecard*.

Walaupun model pada Gambar 9 merupakan model *strategic plan* secara umum, namun perguruan tinggi

juga selayaknya mengadopsinya, karena berdasarkan Karen E. Hinton dari *Society for College and University Planning* (2012) dinyatakan bahwa sebagian lembaga pendidikan tinggi di Amerika Serikat menyusun *masterplan* TI dalam rangka memenuhi kebutuhan *strategic plan* institusi. Hal ini menunjukkan bahwa di berbagai negara maju, IT sudah dipandang sebagai bagian dari *strategic plan* untuk kesuksesan perguruan tinggi (Hinton, 2012).



Gambar 9 Peta Strategi

Jika dilihat pada Gambar 9, dua layer paling bawah mengilustrasikan “*how we plan to accomplish it*”, dan dua layer berikutnya adalah “*what we want to accomplish it*”. Untuk mencapai “*how we plan*” dari sisi IT yaitu melalui *Information Capital*. *Information capital* yang didukung oleh infrastruktur teknologi sebagai salah satu elemen yang akan mendukung keberhasilan *internal operation*, dimana pada akhirnya akan memaksimalkan *organizational value* (Armitage and Scholey, 2006).

Information capital berelasi dengan 5 bagian yaitu : (1) *customer management relationship*, (2) *innovation and commercially supremacy*, (3) *internal operation excellence*, (4) *effective governance and control*, dan (5) *perception (public relation)*.

Teknologi *opinion mining* yang diusung pada penelitian ini adalah untuk mendukung *strategic planning* pada tahap “*how we plan*”. Hasil *opinion mining* yang telah dilakukan berkaitan erat dengan poin *innovation and commercially supremacy* dan *perception (public relation)* serta mendukung *scalability strategies* sehingga memaksimalkan ketercapaian *organization value*.

UIN Sunan Ampel Surabaya sebagai obyek penelitian memiliki visi “menjadi Universitas Islam yang unggul, dan kompetitif bertaraf internasional”. Cita-cita menjadi *World Class University* (WCU) Untuk mendukung tercapainya visi tersebut perlu dilihat *positioning* dari universitas tersebut, salah satunya adalah melalui *webometric*. *Webometric* merilis

kriteria perangkian versi 2018.2.1.3. Dalam laman web *webometric*, metodologi kriteria yang digunakan untuk menilai perangkian web suatu universitas adalah *presence/web visibility*, *impact*, *openness* dan *excellence* (Webometric, 2018)

Gambar 10 menunjukkan hasil pencarian melalui mesin pencari daring bahwa posisi UIN Sunan Ampel menurut *webometric* adalah rangking 4731 dunia, rangking 1788 di Asia, rangking 177 di Asia Tenggara dan rangking 62 di Indonesia (data diambil per 1 November 2018).

| Ranking | World Ranking | University | Det. | Country | Presence | Impact | Openness | Excellence |
|---------|---------------|---|------|---------|----------|--------|----------|------------|
| 62 | 4731 | Universitas Islam Negeri UIN Sunan Ampel Surabaya | | | 936 | 2539 | 5403 | 6008 |

Gambar 10. Posisi UIN Sunan Ampel

Lee menyebutkan bahwa Universitas akan semakin terlihat dari kehadiran/visibilitas dan jejak di Web sebagai pusat reputasi dan kedudukan internasionalnya (Lee and Park, 2012). Dalam konteks ini, web akademik berkembang menjadi lebih dari sekedar wahana untuk mengkomunikasikan prestasi ilmiah dan budaya; konten informasi dipandang sebagai cerminan dari keseluruhan organisasi dan kinerja universitas. Adapun jika dilihat dari sisi *“how we plan”* dalam *strategic planning*, khususnya dalam membangun poin *innovation and commercially supremacy* dan *perception (public relation)* melalui penambangan opini/persepsi di Twitter, UIN Sunan Ampel terlihat jelas kurang memberdayakan *media social* sebagai salah satu media teknologi untuk menunjukkan reputasi dan kedudukannya. Jika dilihat kembali hasil analisis data *Twitter* melalui *word-cloud* (Gambar 5), tidak ada satupun informasi yang memuat publikasi ilmiah, mahasiswa internasional, bahkan kata internasional saja belum ada dalam persepsi masyarakat. Yang terlihat jelas selain nama kampus adalah *“mahasiswa”*, *“nasional”*, *“menpora”*, *“auditorium”*, *“sport center”* dan *“Instagram”*. Kriteria tentang sarana prasarana (*asset utilization*) baru 2 saja yang muncul yaitu auditorium dan *sport center*. Untuk itu, UIN Sunan Ampel perlu memaksimalkan informasi di media sosial, paling tidak memuat informasi yang berhubungan dengan kriteria-kriteria WCU.

Dari segi pengambilan data set selama 2 tahun terakhir, hanya didapatkan 4009 data konten tentang UIN Sunan Ampel. Padahal, setiap 5 menit, orang Indonesia berinteraksi dengan internet terutama media sosial (Hutchison, 2018). Hal ini menunjukkan bahwa UIN Sunan Ampel kurang memberikan informasi, sehingga *follower* yang mengikuti informasi dan menyebarkan informasi tentang UIN Sunan Ampel perlu dikembangkan atau ditambah

volumenya. Sebuah penelitian telah merilis indikator *University Twitter Engagement* untuk merangking institusi berdasarkan pengikut *Twitter* di suatu institusi. Banyaknya *follower* di suatu institusi ternyata erat korelasinya dengan reputasi institusi tersebut (Janssen *et al.*, 2015).

Hasil analisis sentimen dari 2 metode: Vader dan Liu menunjukkan bahwa konten opini yang memuat UIN Sunan Ampel, belum ada yang bernilai kuat positif. Sebagian besar bernilai netral, yang artinya opini/persepsi publik terhadap UINSBY dinilai biasa-biasa saja. Sehingga, UIN Sunan Ampel perlu strategi untuk membangun *innovation and commercially supremacy* serta *perception (public relation)* menjadi lebih kuat nilai positifnya. Gambar 11 menampilkan perbandingan hasil sentimen dari 2 metode yaitu Vader dan Liu. Metode Vader menghasilkan sentiment positif = 2,16%, sentiment negatif = 0,34% dan sentiment netral = 97,54%. Metode Liu menghasilkan nilai sentiment negatif = -50 dan nilai sentiment positif = 50, yang berarti *“netral”*. Sehingga dapat disimpulkan bahwa dengan menggunakan metode Liu maupun Vader memiliki nilai akhir yang sama, yaitu sentimen yang bersifat *“netral”* mendominasi dari 4009 data tweet tentang UIN Sunan Ampel.



Gambar 11. Perbandingan metode Vader dan Liu

Berdasarkan hasil *opinion mining* dari Gambar 11, untuk mendukung *“how we plan” strategic planning* menuju WCU yang merupakan visi misi UIN Sunan Ampel maka perlu dilakukan penguatan *Information Capital* sebagai berikut:

1. Membangun *innovation and commercially supremacy* salah satunya dengan memaksimalkan dan memberdayakan sosial media untuk memperkuat *positioning* di masyarakat melalui peningkatan jumlah *follower*.
2. Menguatkan *perception (public relation)* dengan meningkatkan fungsi media sosial sebagai sarana publikasi yang berkaitan dengan kriteria WCU.

Misal publikasi ilmiah, komunitas keilmuan, diskusi dosen untuk peningkatan akademik, pelayanan pengabdian masyarakat secara daring, dan lain sebagainya.

3. Mempertimbangkan kriteria *University Twitter Engagement* sebagai sarana untuk meningkatkan visibilitas, reputasi dan jejak di dunia maya sebagai bentuk dari *innovation and commercially supremacy*.
4. Meningkatkan *scalability strategies* melalui maksimalisasi fungsi *website* UIN Sunan Ampel.

Terlepas dari *Information Capital*, UIN Sunan Ampel juga perlu mempertimbangkan *Human Capital* dan *Organizational Capital* untuk men-support *Information Capital* yang handal dan berdaya guna.

4. SIMPULAN

Penelitian ini mengambil data opini/persepsi masyarakat tentang UIN Sunan Ampel sebanyak 4009 data tweet. Analisis menggunakan metode Vader menghasilkan sentiment positif = 2,16%, sentiment negatif = 0,34% dan sentiment netral = 97,54%. Sedangkan metode Liu menghasilkan nilai sentiment negatif = -50 dan nilai sentiment positif = 50, yang berarti “netral”. Sehingga dapat disimpulkan bahwa hasil analisis menggunakan metode Liu maupun Vader memiliki nilai akhir yang sama, yaitu sentimen yang bersifat “netral”.

Hasil *opinion mining* tersebut menjadi pijakan dalam mendukung *strategic planning*. *Strategic planning* yang dimaksud adalah unsur *Information Capital* yang didukung oleh infrastruktur teknologi dengan melihat *positioning* kampus di sosial media. Nilai sentimen persepsi masyarakat yang “netral” membuktikan bahwa *Information Capital* tentang UIN Sunan Ampel perlu diperkuat menuju nilai “positif”. Sehingga diperlukan upaya maksimal untuk membangun *innovation and commercially supremacy, perception (public relation)* dan *scalability strategies* supaya *internal operation* bisa handal. Penelitian ini masih jauh dari sempurna, kedepan, untuk merancang dan menghasilkan *masterplan* IT atau *strategic plan* pada tahap “*how we plan*” masih diperlukan riset lanjutan yang membahas tentang *customer management relationship, internal operation excellence, dan effective governance and control* untuk mendukung *Information Capital* dari UIN Sunan Ampel Surabaya.

DAFTAR PUSTAKA

- ARMITAGE, H. M. and SCHOLEY, C. (2006) *Management Accounting Guideline, Using Strategy Maps to Drive Performance*. Available at: http://www.cimaglobal.com/Documents/ImportedDocuments/Tech_MAG_Strategy_Mapping_March07.pdf.
- AUINGER, A. and FISCHER, M. (2008) ‘Mining consumers’ opinions on the web’, *FH Science Day*, (Linz, Österreich), pp. 410–419.
- DEVIKA, M. D., SUNITHA, C. and GANESH, A. (2016) ‘Sentiment Analysis: A Comparative Study on Different Approaches’, *Procedia Computer Science*. The Author(s), 87, pp. 44–49. doi: 10.1016/j.procs.2016.05.124.
- FANG, X. and ZHAN, J. (2015) ‘Sentiment analysis using product review data’, *Journal of Big Data*. *Journal of Big Data*, 2(1). doi: 10.1186/s40537-015-0015-2.
- HADDI, E., LIU, X. and SHI, Y. (2013) ‘The role of text pre-processing in sentiment analysis’, *Procedia Computer Science*. Elsevier B.V., 17, pp. 26–32. doi: 10.1016/j.procs.2013.05.005.
- HAFNIKA, F. *et al.* (2016) ‘Strategy Map Formulation For Designing Strategic Plan’, *Advance Business and Social Studies*, 2(2), pp. 103–112. Available at: https://apiar.org.au/wp-content/uploads/2016/08/8-APJABSS_ICABSS-BRR732_BIZ-103-112.pdf.
- HINTON, K. E. (2012) *A Practical Guide to Strategic Planning in Higher Education*, *Society for College and University Planning*. doi: 10.1073/pnas.1719902115.
- HUTCHISON, C. S. (2018) ‘Essential Insight Into Internet, Social Media, Mobile, and Ecommerce Use Around the World’, *In: the Ocean basins and margins: The Indian Ocean (Nairn, A.E.M. & Stehli, F.G., Eds) Plenum Press, New York*, 6, pp. 451–512.
- HUTTO, C. J. and GILBERT, E. (2014) ‘VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text’, *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, pp. 216–225.
- JANSSEN, D. *et al.* (2015) ‘Sentiment Analysis of Social Media for Evaluating Universities’, pp. 49–62.
- LEE, M. and PARK, H. W. (2012) ‘Exploring the web visibility of world-class universities’, *Scientometrics*, 90(1), pp. 201–218. doi: 10.1007/s11192-011-0515-6.
- LI, Q. *et al.* (2018) ‘Web Media and Stock Markets: A Survey and Future Directions from a Big Data Perspective’, *IEEE Transactions on Knowledge and Data Engineering*, 30(2), pp. 381–399. doi: 10.1109/TKDE.2017.2763144.

- LV, X. and EL-GOHARY, N. (2016) 'Text Analytics for Supporting Stakeholder Opinion Mining for Large-scale Highway Projects', *Procedia Engineering*. Elsevier B.V., 145, pp. 518–524. doi: 10.1016/j.proeng.2016.04.039.
- MCCOY, C. G., NELSON, M. L. and WEIGLE, M. C. (2017) 'University Twitter Engagement: Using Twitter Followers to Rank Universities'. Available at: <http://arxiv.org/abs/1708.05790>.
- MEDHAT, W., HASSAN, A. AND KORASHY, H. (2014) 'Sentiment analysis algorithms and applications: A survey', *Ain Shams Engineering Journal*. Faculty of Engineering, Ain Shams University, 5(4), pp. 1093–1113. doi: 10.1016/j.asej.2014.04.011.
- MISHRA, N. (2012) 'Classification of Opinion Mining Techniques', 56(13), pp. 1–6.
- 'Orange3 Text Mining Documentation' (2018).
- PANG, B. and LEE, L. (2006) 'Opinion Mining and Sentiment Analysis', *Foundations and Trends® in Information Retrieval*, 1(2), 91–231. doi:10.1561/1500000001n Retrieval, 1(2), pp. 91–231. doi: 10.1561/1500000001.
- PUTRANTI, N. D. and WINARKO, E. (2014) 'Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine', *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 8(1), pp. 91–100. doi: 10.22146/ijccs.3499.
- RITSON, N. (2011) *The Levels and Formulation of Strategy, Strategic Management*.
- ROZI, I. F., PRAMONO, S. H. and DAHLAN, E. A. (2012) 'Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi', *Electrical Power, Electronics, Communications, Controls, and Informatics Seminar (EECCIS)*, 6(1), pp. 37–43.
- SCIENCE, C. and ENGINEERING, S. (2016) 'Algorithms for Opinion Mining and Sentiment Analysis : An Overview', 6(2), pp. 455–459.
- SINGH, P. K. and HUSAIN, M. S. (2014) 'Methodological Study of Opinion Mining and Sentiment Analysis Techniques', 5(1), pp. 11–21. doi: 10.5121/ijsc.2014.5102.
- SINGLA, Z., RANDHAWA, S. and JAIN, S. (2017) 'Sentiment Analysis of Customer Product Reviews Using Machine Learning'.
- STEFANOVSKA, L. and SOKLEVSKI, T. (2014) 'Benefits of Using Balanced Scorecard in Strategic and Operational Planning', *Universal Journal of Management*, 2(4), pp. 165–171. doi: 10.13189/.
- STIEGLITZ, S. *et al.* (2014) 'Socialmedia analytics', *Business and Information Systems Engineering*, 6(2), pp. 89–96. doi: 10.1007/s12599-014-0315-7.
- STIEGLITZ, S. *et al.* (2018) 'Social media analytics – Challenges in topic discovery, data collection, and data preparation', *International Journal of Information Management*. Elsevier, 39(October 2017), pp. 156–168. doi: 10.1016/j.ijinfomgt.2017.12.002.
- TARTIR, S. and ABDUL-NABI, I. (2017) 'Semantic Sentiment Analysis in Arabic Social Media', *Journal of King Saud University - Computer and Information Sciences*. King Saud University, 29(2), pp. 229–233. doi: 10.1016/j.jksuci.2016.11.011.
- THAMRIN, H. and PAMUNGKAS, E. W. (2017) 'A Rule Based SWOT Analysis Application: A Case Study for Indonesian Higher Education Institution', *Procedia Computer Science*. Elsevier B.V., 116(December), pp. 144–150. doi: 10.1016/j.procs.2017.10.056.
- WEBOMETRIC (2018) *The Ranking Web or Webometrics Methodology*. Available at: <http://www.webometrics.info/en/Methodology>.