

## PEMETAAN RISET TENTANG DETEKSI TOPIK PADA TWITTER DENGAN TEKNIK SYSTEMATIC LITERATURE REVIEW

Sri Suhartini<sup>1</sup>, Widodo<sup>2</sup>, Bambang Prasetya Adhi<sup>3</sup>

<sup>1</sup> Mahasiswa Prodi Pend. Teknik Informatika dan Komputer, Teknik Elektro, FT – UNJ

<sup>2,3</sup> Dosen Prodi Pend. Teknik Informatika dan Komputer, Teknik Elektro, FT – UNJ

<sup>1</sup>[suhartinisri.17@gmail.com](mailto:suhartinisri.17@gmail.com), <sup>2</sup>[widodo@unj.ac.id](mailto:widodo@unj.ac.id), <sup>3</sup>[bambang\\_padhi@yahoo.com](mailto:bambang_padhi@yahoo.com)

### Abstrak

Twitter merupakan salah satu media sosial yang banyak digemari orang, mulai dari anak remaja hingga dewasa. Di twitter banyak informasi yang akan didapatkan oleh para pengguna. Masing-masing informasi memiliki kata kunci untuk dikelompokkan topik bahasannya. Kata kunci yang dimaksud adalah kata yang memiliki posisi yang istimewa dalam sebuah kalimat. Dari kata kunci ini maka topik akan lebih mudah untuk di deteksi. Terdapat berbagai macam metode untuk mendeteksi topik pada twitter. Penelitian ini bertujuan untuk membuat peta riset deteksi topik pada twitter dengan teknik *systematic literature review*. Literatur yang digunakan berasal dari beberapa digital library yaitu IEEE eXplore, Scencedirect, Google Scholar dan Citeseerx. Selanjutnya proses penyaringan literatur berdasarkan relevansi judul didapat 50 judul literatur. Setelah itu menganalisis isi dari masing-masing literatur. Hasil analisis menunjukkan terdapat 39 jenis metode deteksi topik. Metode terbanyak yang digunakan ialah Latent Dirichlet Allocation (LDA). Dan metode dengan tingkat akurasi tertinggi yang digunakan ialah Random Forest (RF) dan Naives Bayes dengan tingkat akurasi 95,5%.

**Kata kunci :** Deteksi Topik, *Systematic Literature Review*, Twitter.

### 1. Pendahuluan

Salah satu faktor penting penunjang globalisasi ialah internet. Internet banyak dimanfaatkan untuk berbagai pemenuhan kebutuhan, dimana salah satunya dimanfaatkan untuk sosial media. Saat ini banyak sosial media yang digemari oleh masyarakat, baik dari luar negeri maupun dalam negeri. Salah satu contoh sosial media yang sedang *trend* saat ini yaitu *twitter*.

*Twitter* banyak digunakan berbagai kalangan masyarakat, mulai dari remaja hingga orang dewasa. Yang membuat *twitter* ini unik ialah karena *twitter* menerapkan teks pendek pada penggunaannya. Pada *Twitter*, *user* hanya dapat menulis status atau yang dikenal dengan *tweet* sebanyak 280 karakter.

Salah satu topik yang dibahas mengenai *twitter* ialah metode deteksi topik pada *twitter*. *Twitter* mempunyai banyak fitur di dalam aplikasinya, salah satunya adalah *trending topic*. *Trending topic* adalah suatu hal yang sedang terkenal dan banyak dibicarakan dalam kurun waktu tertentu (Suryadi: 2018). *Trending topic* berisi dari 5 sampai 20 daftar topik terpopuler yang sedang berlangsung. Topik-topik tersebut didaftar berdasarkan lokasi negara *user* itu sendiri. *Trending topic* yang terdaftar diperingkat berdasarkan banyaknya *tweet* yang

membahas topik tersebut. Maka setiap negara akan memiliki *trending topic* yang berbeda.

Pada *twitter* terdapat kesulitan dalam pendeteksian topik. Hal ini karena *twitter* menerapkan sistem *real time* dalam mendeteksi topik dari *tweet* yang ditulis oleh *user*. Sistem *real time* ini yaitu mendeteksi topik secara langsung pada saat *tweet* itu dipublikasikan. Beberapa riset telah dilakukan dengan berbagai macam metode dalam pendeteksian topik pada *twitter*. Namun hingga kini, berdasarkan publikasinya belum ada pemetaan riset mengenai deteksi topik pada *twitter*. Maka, perlu dibuat sebuah model pemetaan dalam pendeteksian topik pada *twitter*. Dengan adanya pemetaan ini, maka akan didapatkan beberapa poin penting, seperti jenis-jenis metode yang digunakan dalam pendeteksian topik pada *twitter*, metode terbanyak yang digunakan dalam pendeteksian topik pada *twitter*, dan metode dengan tingkat akurasi tertinggi dalam pendeteksian topik pada *twitter*.

Untuk membuat pemetaan riset deteksi topik pada *twitter* ini digunakan metode *systematic literature review* (SLR). Penggunaan metode ini dianggap paling efektif dalam memetakan riset, sebab dengan metode SLR data riset yang dipakai akan diekstrak dan dianalisis secara mendalam, sehingga menghasilkan keluaran data dengan

Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>

kualitas yang baik dan dapat dipertanggungjawabkan.

## 2. Dasar Teori

### 2.1 Systematic Literature Review

*Systematic literature review* atau sering disingkat SLR adalah metode *literature review* yang mengidentifikasi, menilai, dan menginterpretasi seluruh temuan-temuan pada suatu topik penelitian, untuk menjawab pertanyaan penelitian (*research question*) yang telah ditetapkan sebelumnya (Kitchenham & Charters, 2007). Metode SLR dilakukan secara sistematis dengan mengikuti tahapan dan protokol yang memungkinkan proses *literature review* terhindar dari bias dan pemahaman yang bersifat subyektif dari penelitiannya. Pada saat melakukan *review* menggunakan metode SLR diharuskan melalui beberapa tahapan yaitu *planning, conducting, dan reporting*

### 2.2 Pemetaan Riset

Pemetaan riset merupakan alat visualisasi yang mengatur dan menampilkan pengetahuan dengan cara visual, seperti menggunakan tabel, diagram alir, grafik dan diagram venn. Lingkup tinjauan yang digunakan juga melacak studi dengan ulasan sistematis. Dibanding dengan tinjauan sistematis, lingkup tinjauan lebih difokuskan pada temuan awal, meskipun tidak selalu memiliki kualitas sebagai data awal untuk dimasukkan, dan sintesis data lebih terfokus pada kualitatif (Diane Cooper : 2016).

### 2.3 Deteksi Topik

Tugas deteksi topik ialah mengevaluasi teknologi yang mendeteksi topik-topik baru yang sebelumnya tidak dikenal. Seperti dalam tugas pelacakan, topik didefinisikan dengan mengaitkan cerita yang membahas topik tersebut. Namun, sistem pendeteksian topik tidak diberi pengetahuan tentang topik tersebut (Jonathan : 2018).

### 2.4 Twitter

*Twitter* adalah layanan jejaring sosial dan mikroblog daring yang memungkinkan penggunaanya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter, akan tetapi pada tanggal 7 November 2017 bertambah hingga 280 karakter yang dikenal dengan sebutan kicauan atau tweet (Jack Dorsey : 2018).

## 3. Metodologi

### 3.1 Tempat dan Waktu Penelitian

Penelitian ini dilakukan di Laboratorium Komputer Gedung L1 Teknik Elektro lantai 2, Fakultas Teknik, Universitas Negeri Jakarta yang berlokasi di Jalan. Rawamangun Muka, Jakarta 13220. Penelitian ini dilakukan sejak bulan Januari 2018 hingga Juli 2019.

## 3.2 Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah metode *systematic literature review*. Dimana peneliti melakukan analisis pada penelitian yang telah dilakukan sebelumnya melalui literatur-literatur penelitian yang berkaitan dengan klasifikasi teks pendek pada *twitter*.

## 3.3 Data dan Sumber Data

Data yang digunakan berupa dokumen hasil penelitian dalam bentuk literatur yang didapatkan dengan cara melakukan proses mengunduh pada *digital library* terpercaya. Literatur yang digunakan merupakan hasil penelitian sebelumnya yang relevan dengan apa yang dibutuhkan penulis dalam penelitian ini.

## 3.4 Teknik dan Prosedur Pengumpulan Data

Pengumpulan data untuk penelitian ini menggunakan data dari *digital library* IEEE *eXplore*, *Sciencedirect*, *Google Scholar* dan *Citeseerx* dalam pencarian literatur terkait. Alasan dari pemilihan ke empat *digital library* tersebut dikarenakan merupakan *digital library* yang tersedia di Universitas Negeri Jakarta dan juga tidak adanya kriteria atau karakteristik khusus dalam pemilihan *digital library*.

## 3.5 Prosedur Analisis Data

Pada penelitian ini, dibuat sebuah aturan yang mengikat yaitu literatur yang membahas metode deteksi topik pada *twitter*, literatur dengan batas tahun publikasi 2014-2018, dan literatur berbahasa inggris dan pada saat pencarian di *digital library* menggunakan *search string: (topic detection) AND (twitter)*.

## 3.6 Pemeriksaan Keabsahan Data

Pemeriksaan keabsahan data dilakukan berdasarkan relevansi judul dari literatur yang tersedia dengan topik bahasan yaitu deteksi topik pada *twitter*, isi dari abstrak dan diskusi dengan pihak yang memiliki kompetensi dan berkepentingan dalam penelitian riset deteksi topik pada *twitter* menggunakan metode *systematic literature review*. Dalam hal ini yang dimaksud adalah dosen pembimbing.

## 4. Hasil dan Analisis

### 4.1 Planning

Pada tahapan *planning* dimulai dengan merumuskan masalah yang akan diteliti. Pada penelitian ini, peneliti mengangkat masalah mengenai metode yang digunakan untuk pendeteksi topik pada *twitter*. Setelah perumusan masalah, maka dilakukan identifikasi masalah dan bagian terakhir dari tahap *planning* ialah menentukan *Research Question*.

*Research Question* dibuat secara spesifik untuk membuat *review* menjadi terfokus, untuk

Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>

merumuskan *research question* yang efektif diperlukan ketentuan *Population, Intervention, Comparison, Outcomes* dan *Context* (PICOC) yang merupakan lima faktor utama dalam menentukan *research question* sehingga mempermudah peneliti dalam perumusan RQ. Format dari PICOC sendiri tidak memiliki kriteria tertentu, sehingga peneliti bebas menentukan format seperti apa yang paling sesuai dengan peneliti. Pada penelitian ini, format tabel yang digunakan peneliti dapat dilihat di tabel 4.1

**Tabel 4.1 Faktor PICOC**

FAKTOR	
<i>Population</i>	<i>Twitter</i>
<i>Intervention</i>	Metode deteksi topik
<i>Comparison</i>	Metode deteksi topik dan tingkat keberhasilan
<i>Outcomes</i>	Metode deteksi topik dan tingkat keberhasilan
<i>Context</i>	Akurasi metode deteksi topik

Dari kelima faktor PICOC diatas maka dirumuskanlah *research question* yang dapat dilihat pada tabel 4.2.

**Tabel 4.2 Research Question**

ID	<i>Research Question</i>
RQ1	Apa jenis – jenis metode yang digunakan dalam riset pendeteksian topik pada <i>twitter</i> ?
RQ2	Apa metode yang paling sering digunakan dalam riset pendeteksian topik pada <i>twitter</i> ?
RQ3	Apa metode dengan tingkat akurasi tertinggi yang digunakan dalam riset pendeteksian topik pada <i>twitter</i> ?
RQ4	Asal dari sumber data dan berapa banyak data yang digunakan dalam riset pendeteksian topik pada <i>twitter</i> ?
RQ5	Siapa penulis yang paling sering melakukan riset pendeteksian topik pada <i>twitter</i> ?
RQ6	Jurnal dan <i>conference</i> apa yang paling sering menjadi <i>event</i> untuk topik deteksi ?

## 4.2 Conducting

Pada tahap *conducting*, yang pertama kali dilakukan ialah mencari atau mengidentifikasi penelitian yang relevan dengan penelitian yang akan dilakukan. Pencarian atau identifikasi pada *digital library* dihasilkan dari memasukkan *key word* pada mesin pencarian yang sebelumnya telah ditentukan. Penggunaan *key word* dilakukan secara konsisten pada setiap pencarian literatur pada mesin pencarian *digital library* untuk memaksimalkan pencarian literature yang relevan.

Dari hasil pencarian tersebut didapat sekitar 93 literatur berbahasa inggris dengan topik bahasan “deteksi topik pada *twitter*”. Dari 93 literatur yang ada, 18 didapat dari *IEEE eXplore*, 11 didapat dari *Citeseerx*, 20 didapat dari *Google Scholar*, dan 44 didapat dari *Science Direct*.

Dari 93 literatur yang ada, berdasarkan hasil identifikasi, literatur dipilih kembali berdasarkan relevansi isi dalam penelitian yang dilakukan. Relevansi isi ini disaring berdasarkan pemenuhan kriteria. Kriteria yang dipakai ialah membahas mengenai metode deteksi topik pada *twitter*. Untuk memastikan kualitas setiap literatur, maka peneliti perlu membaca keseluruhan 93 literatur. Setelah melakukan identifikasi lanjutan, dengan adanya penyaringan abstrak dan isi, maka literatur yang memiliki relevansi isi berubah menjadi 50 literatur.

Setelah melakukan pengkajian atau analisis data, hasil dari analisis data tersebut di masukkan ke dalam *form* ekstraksi data yang dibuat berdasarkan *research question* yang telah dirumuskan sebelumnya. *Form* ekstraksi data dapat dilihat pada lampiran 1.

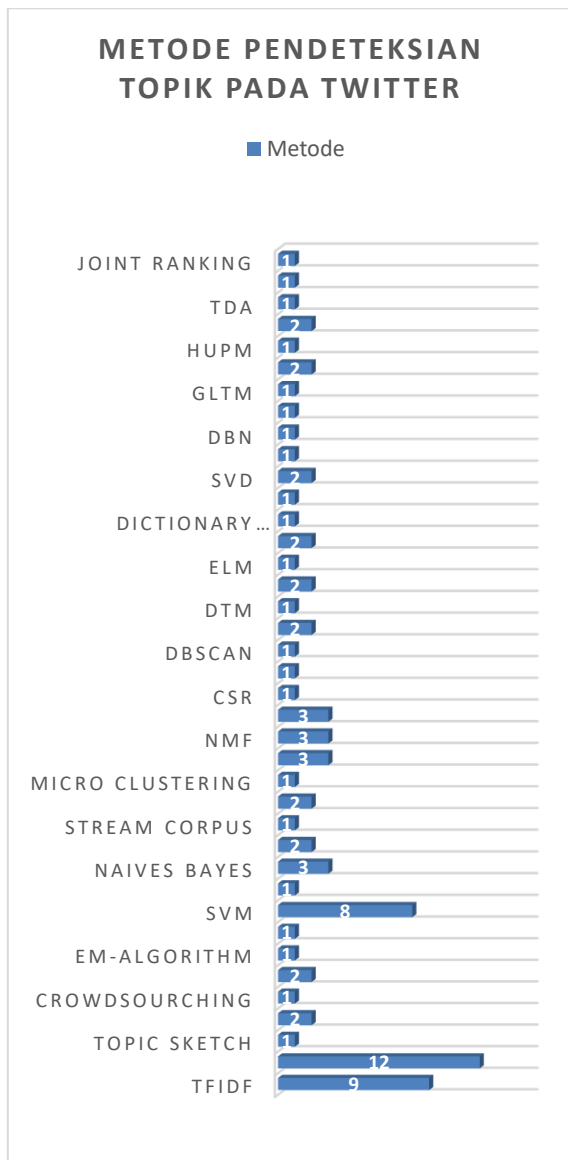
Setelah mendata hasil dari 50 literatur yang ada ke dalam *form* ekstraksi, maka dilakukanlah sintesis data untuk menjawab *research question* yang telah dibuat sebelumnya. *Research question* yang utama merupakan RQ1, RQ2, dan RQ3. Sedangkan RQ4, RQ5, dan RQ6 merupakan *research question* tambahan.

### 4.2.1 Metode Deteksi Topik pada *Twitter*

Dari gambar 4.1 dapat dilihat bahwa terdapat 39 macam metode deteksi topik yang digunakan dalam pendeteksian topik pada *twitter*.

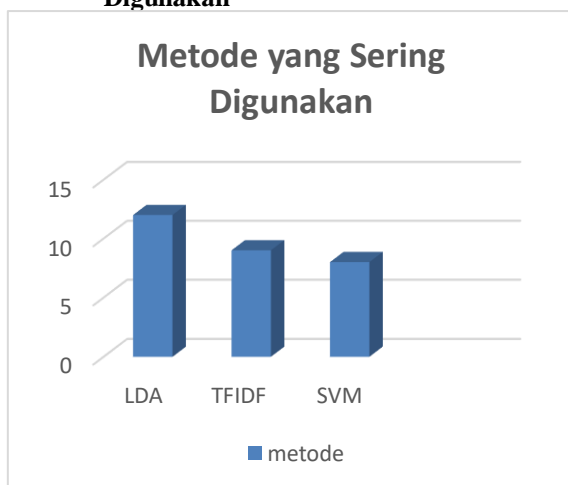
Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>



Gambar 4.1 Macam-macam Metode Deteksi Topik pada Twitter

#### 4.2.2 Metode Deteksi Topik yang Sering Digunakan



Gambar 4.2 Metode Paling Sering Digunakan

Dapat dilihat pada gambar 4.2 bahwa tiga jenis metode deteksi topik teratas yang banyak digunakan ialah *Latent Dirichlet Allocation* (LDA) sebanyak 12 judul, *Term Frequency – Inverse Document Frequency* (TFIDF) sebanyak 9 judul, dan *Support Vector Machine* (SVM) sebanyak 8 judul. Dari ketiga jenis yang banyak digunakan ini, terdapat satu jenis metode deteksi topik yang paling sering digunakan, yaitu LDA. Penggunaan metode LDA dinilai efektif dan dapat meminimalisir suatu kesalahan dalam mendeteksi topik pada *twitter*. Metode LDA sendiri sudah banyak teori dan penelitian terkait yang tersedia secara umum, sehingga memudahkan para peneliti dalam menggunakannya.

#### 4.3 Metode Klasifikasi Terbaik

Berdasarkan tingkat akurasinya penelitian berjudul *A holistic system for troll detection on twitter* menggunakan metode *Random Forest* (RF) dan *Naives Bayes* menempati posisi tertinggi dengan nilai akurasi sebesar 95,5%. Nilai akurasi tertinggi lainnya diikuti oleh judul *Exemplar-based topic detection in twitter streams* menggunakan metode SVD, TFIDF, dan NMF dengan nilai akurasi 95%. Nilai akurasi tertinggi ketiga ditempati oleh judul *Detecting Malicious Tweets in Trending Topics using Clustering and Classification* menggunakan metode K-Means dan ELM dengan tingkat akurasi sebesar 92%.

#### 4.4 Sumber dan Banyak Data dalam Riset

Dalam riset pendeteksian topik pada *twitter* para penulis mendapatkan data nyata dari *Application Programming Interface* (API) yang disediakan oleh pihak *Twitter* guna untuk mendapatkan informasi berupa data diri, *tweet*, *mention*, *retweet*, *favourit*, *followers*, dan teman dari akun *twitter* tertentu yang terhubung dengan aplikasi. Rentang waktu pengambilan data beragam, mulai dari 24 jam, 20 hari, hingga 3 bulan. Banyak data yang dipakai juga berbeda – beda sesuai dengan lamanya pengambilan data, mulai dari 200 *tweets*, 1.500 *tweets*, 12.000 *tweets*, dan 57.414 *tweets*. Banyak dan lamanya data yang digunakan tergantung pada penelitian yang dilakukan.

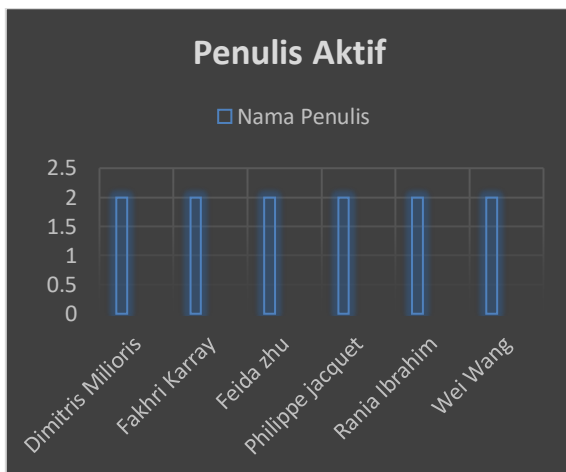
#### 4.5 Penulis Aktif

Dari 50 literatur yang diambil, terdapat 171 *researcher* yang berkontribusi pada penelitian terkait pendeteksian topik pada *twitter*, dengan detail 165 diantaranya berkontribusi pada 1 penelitian dan 6 lainnya berkontribusi pada 2 penelitian. 6 peneliti ini ialah Dimitris Milioris, Fakhri Karray, Feida Zhu, Phillipe Jacquet, Rania Ibrahim dan Wei Wang. Penulis aktif yang banyak melakukan penelitian terkait merupakan penulis pertama dan bukan penulis pertama.

Available at:

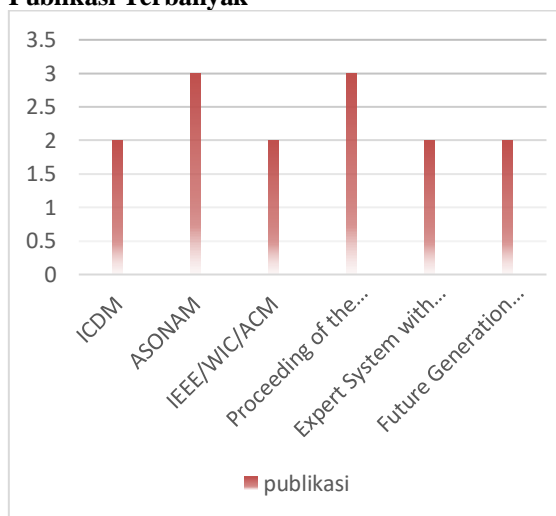
<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>





Gambar 4.3 Topik Riset Analisis pada Twitter

#### 4.6 Publikasi Terbanyak



Gambar 4.4 Publikasi Terbanyak

Tercatat dalam kurun waktu 5 tahun, dari tahun 2014 hingga 2018 terdapat banyak *event* yang mempublikasikan riset mengenai deteksi topik pada *twitter*. *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (ASONAM) mempublikasikan 3 judul penelitian dan *Proceedings of the SNOW Data Challenge* mempublikasikan 3 judul penelitian..

#### 4.7 Pembahasan

Dari 50 literatur, didapatkan 39 jenis metode deteksi topik yaitu *Term Frequency-Inverse Document Frequency* (TF-IDF), *Latent Dirichlet Allocation* (LDA), *Topic Sketch*, *User Temporal Mixture* (UTM), *Crowdsourcing*, *Time User Sentiment LDA* (TUS-LDA), *Expectation Maximization* (EM), *Rumour Detection System*, *Support Vector Machine* (SVM), *Mixture Model Approach*, *Naives bayes*, *Random Forest* (RF), *Stream Corpus*, *On-line LDA* (OLDA), *Micro Clustering*, *Burst Method*, *Non-negative Matrix Factorization* (NMF), *Topic Graph*, *Concept Space*

*Representation* (CSR), *Semantic Space Representation* (SSR), *Density Based Spatial Clustering* (DBSCAN), *Joint Complexity*, *Dynamic Topic Model* (DTM), *K-Means Clustering*, *Extreme Learning Machine* (ELM), *Document Pivot*, *Dictionary Learning*, *Real-Time Framework* (RTF), *Singular Value Decomposition* (SVD), *Formal Concept Analysis* (FCA), *Deep Belief Network* (DBN), *Real-Time non Parametric* (RTP), *Global and Local Topic Model* (GLTM), *Spam Post Detection*, *High Utility Pattern Mining* (HUPM), *BN-Grams*, *Topic Detection using AGF* (TDA), *FP-Streams*, dan *Joint Ranking*.

Dari ke-39 jenis metode deteksi topik yang ada, terdapat tiga metode yang paling banyak digunakan. Posisi pertama ialah metode *Latent Dirichlet Allocation* (LDA) yang digunakan oleh 12 judul penelitian. Penggunaan metode LDA dinilai efektif dan dapat meminimalisir suatu kesalahan dalam mendeteksi topik pada *twitter*. Metode LDA sendiri sudah banyak teori dan penelitian terkait yang tersedia secara umum, sehingga memudahkan para peneliti dalam menggunakannya. Posisi kedua terdapat metode *Term Frequency-Inverse Document Frequency* (TF-IDF) yang digunakan oleh 9 judul penelitian dan ketiga terdapat *Support Vector Machine* (SVM) yang digunakan oleh 8 judul penelitian.

Dari 50 judul literatur yang ada, berdasarkan tingkat akurasi penelitian berjudul *A holistic system for troll detection on twitter* dengan menggunakan metode *Random Forest* (RF) dan *Naives Bayes* menempati posisi tertinggi dengan nilai akurasi sebesar 95,5%. Dalam penelitian ini telah diidentifikasi 6 kelompok fitur, masing-masing berdasarkan analisis gaya penulisan, sentiment, perilaku, interaksi sosial, media terkait, dan waktu publikasi. Hasil penelitian menunjukkan bahwa klarifikasi otomatis dapat berguna dalam seluruh proses identifikasi dan pengelolaan perilaku anti-sosial online yang terdapat pada aplikasi *twitter*. Nilai akurasi tertinggi lainnya diikuti oleh judul *Exemplar-based topic detection in twitter streams* yang menggunakan metode SVD, TFIDF, dan NMF dengan nilai akurasi 95%. Nilai akurasi tertinggi ketiga ditempati oleh judul *Detecting Malicious Tweets in Trending Topics using Clustering and Classification* yang menggunakan metode K-Means dan ELM dengan tingkat akurasi sebesar 92%.

## 5 Kesimpulan dan Saran

### 5.1 Kesimpulan

Kesimpulan yang diperoleh dari pemetaan riset tentang deteksi topik pada *twitter* setelah melalui seluruh tahapan *systematic literature review* dapat dirincikan dari hasil penelitian sebagai berikut ini:

- 1) Dari 50 literatur yang ada terdapat 39 jenis metode pendeteksian topik pada *twitter*.

Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>

- 2) Metode yang paling banyak digunakan untuk pendeteksian topik pada *twitter* ialah *Lantent Dirichlet Allocation* (LDA) sebanyak 12 judul lalu diikuti oleh *Term Frequency – Inverse Document Frequency* (TF-IDF) sebanyak 9 judul dan *Support Vector Machine* (SVM) sebanyak 8 judul. LDA dinilai efektif dan dapat meminimalisir suatu kesalahan dalam mendeteksi topik pada *twitter*.
- 3) Metode dengan tingkat akurasi tertinggi yang digunakan ialah *Random Forest* (RF) dan *Naives Bayes* dengan nilai akurasi 95,5%. Diikuti oleh metode SVD, TFIDF, dan NMF dengan nilai akurasi 95% serta metode K-Means dan ELM dengan nilai akurasi 92%.
- 4) Sumber data yang dipakai dalam penelitian didapat dari *Application Programming Interface* (API). Banyak data yang digunakan beragam, mulai dari 200 *tweets*, 1.500 *tweets*, 12.000 *tweets*, dan 57.414 *tweets*.
- 5) Terdapat 6 peneliti aktif dalam penelitian ini yaitu Dimitris Milioris, Fakhri Karray, Feida Zhu, Phillipe Jacquet, Rania Ibrahim dan Wei Wang yang masing - masing berpartisipasi dalam 2 judul penelitian.
- 6) Publikasi terbanyak terdapat dalam *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (ASONAM) dan *Proceedings of the SNOW Data Challenge* yang masing – masing mempublikasi 3 judul penelitian.

## 5.2 Saran

Penulis memiliki masukan untuk penelitian selanjutnya yang berhubungan dengan metode deteksi topik dengan teknik *systematic literature review* sebagai berikut:

- 1) Menggunakan lebih banyak *digital library* untuk mendapatkan literature, agar data yang diperoleh dapat maksimal.
- 2) Menambahkan *research question* agar pemetaan yang dibuat lebih mendetail.

## Daftar Pustaka:

- Abbas, Rini. Tanpa Tahun. PEMETAAN\_2, [https://www.academia.edu/16537674/PEMETAAN\\_2](https://www.academia.edu/16537674/PEMETAAN_2). Diakses 30 April 2018.
- A.G., Haryanto. 1999. Metode Penulisan dan Penyajian Karya Ilmiah: Buku Ajar untuk Mahasiswa. Jakarta: Buku Kedokteran EGC.
- Benny, Amrutha., Philip, Mintu. 2015. Keyword based tweet extraction and detection of related topics. 2015, *Procedia Computer Science volume 46*.
- Burnside, Gerard., Milioris, Dimitris., Jacquet, philippe. 2014. One day in twitter: topic detection via joint complexity. 2014 *In Proceedings of the SNOW Data Challenge, Seoul, Korea*
- Cataldi, Mario., Caro, L. D., Schifanella, Claudio. 2014. Emerging topic detection on twitter based on temporal and social terms evaluation. 2014, *Proceeding of the Tenth International Workshop on Multimedia Data Mining*.
- Chen, Yan., Li, Zhoujun., Chua, Tat-Seng. 2014. Emerging topic detection for organizations from microblogs. 2014 *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*.
- Choi, Hyeok-Jun., Park, C.H., Emerging topic detection in twitter stream based on high utility pattern mining. 2018, *Expert Systems with Applications 115*.
- Cigarran, Juan., Castellanos, Angel., Garcia, Ana. 2016. A step forward for topic detection in twitter an FCA based-approach. *Expert Syst. Appl. 2016*
- Cooper, Diane. 2016. What is a “mapping study?”. *Journal of the Medical Library Association*.
- Dang, Qi., Gao, Feng., Zhou Yadong. 2015. Early detection method for emerging topics based on dynamic bayesian networks in micro-blogging networks. 2015 *Current Trends in Web Engineering: 15th International Conference, ICWE 2015 Workshops*.
- Djunaedi, achmad. 2000. Pengertian masalah, variabel, dan paradigma penelitian <https://farelbae.wordpress.com/catatan-kuliah-ku/pengertian-masalah-variabel-paradigma-penelitian/>. Diakses 6 desember 2018.
- Dong, Guozhong., Yang, Wu., Zu, Feida., Wang, Wei. 2016. Discovering burst patterns of burst topic in twitter. 2016, *Computer and electrical Engineering 58*.
- Dorsey, Jack. 2018. Just Setting Up My Twitter. *Wikipedia*. Diakses 10 Januari 2019.
- Eitel, Barry. 2018. Twitter ubah algoritma untuk kurangi trolling. <https://www.aa.com.tr/id/dunia/twitter-ubah-algoritma-untuk-kurangi-trolling/1147473>. Diakses 30 Desember 2018.
- Elbagoury, Ahmed., Ibrahim, Rania., Farahat, K. A., Kamel, M. S., Karray, Fakhri. 2015. Exemplar-based topic detection in twitter streams. 2015 *International AAAI Conference on Web and Social Media*.
- Fornacciari, Paolo., Mordonini, Monica., Poggi, Agostino., Sani, Laura., Tomaiuolo, M. 2018. A holistic system for troll detection on twitter. 2018, *Computers in Human Behavior*.
- Guille, Adrien., Favre, Cecile. 2014. Mention-anomaly-based events detection and

Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>

- tracking in twitter. *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*.
- Guo, Jing., Zhang, Peng., Tan, Jianlog., Guo, Li. 2014. Mining hot topics for twitter streams. *2014, Procedia Computer Science volume 9*.
- Hall, T., Beecham S., Bowes D., Gray D., & Counsell, S. (2012). A Systematic Literature Review on Fault Prediction Performance in Software Engineering. *IEEE Transactions on Software Engineering*, 38:1276-1304.
- Hagras, Mohamed., Hassan, Ghada., Farag, Nadine. 2017. Towards natural disasters detection from twitter using topic modelling. *2017 European Conference on Electrical Engineering and Computer Science (EECS)*.
- Hammad, Muhammad., Elbeltagy, Samhaar. 2017. Towards efficient online topic detection through automated bursty feature detection from arabic twitter streams. *Procedia Computer Science 117:248-255 · November 2017*.
- Haryanto. 2000. Metode Penulisan dan Penyajian Karya Ilmiah vol. 1. Jakarta: EGC.
- Hayashi, Kohei., Maehara, Takatori., Toyoda, Masashi, Kawarabayashi, K. 2015. Real-time top-R topic detection on twitter with topic hijack filtering. *2015 the 21th ACM SIGKDD International Conference*.
- Heck, P. & Zaidman, A. (2016). A Systematic Literature Review on Quality Criteria for Agile Requirements Specifications. *Link Springer*, 26:127-160.
- Huang, Siqi., Yang Y., Huakang L., Guoozi S. 2014. Topic Detection from Microblog Based on Text Clustering and Topic Model Analysis. *2014 Asia-Pacific Services Computing Conference*.
- Huang, Bo., Yang, Yan., Mahmood, A., Wang, Hongjum. 2014. Microblog topic detection based on LDA model and single-pass clustering. *8th international conference, RSCTC 2014, Chengdu, China, Proceedings*.
- Indra., Winarko, Edi., Pulungan, Reza. 2018. Trending topics detetection of indonesian tweets using BN-grams and doc-p. *2018, Computer and Information Science Vol. 31 Issue 2*.
- Inuwa-dutse, Isa., Liptrott, Mark., Korkontzelos, Ioannis. 2018. Detection of spam-posting accounts on twitter. *2018, Neurocomputing 315*.
- Ishikawa, Shota., Arakawa, Yukata., Tagashira, Shigeaki., Fukuda, Akira. 2014. Hot topic detection in local areas using twitter. *ARCS 2014*.
- Kim, Hwi-Gang., Lee, S., Kyeong, S. 2014. Discovering Hot Topics Using Twitter Streaming Data Social Topic Detection and Geographic Clustering. *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM )*.
- Kitchenham & Charters. (2007). Guidelines in Performing Systematic Literature Reviews in Software Engineering. *EBSE Technical Report 33(5) : 20*.
- Koike, daichi., Yoshioka, Masaharu., Kando, Noriko. 2014. Time series topic modelling and bursty topic detection of correlated news and twitter. *IJCNLP 2014*.
- Lau, Ha. J., Collier, Nigel., Baldwin, Timothy. 2014. On-line trend analysis with topic models: twitter trends detection topic model online. *Proceedings of COLING 2014*.
- Liesnoor, Dewi dkk .2001. Desain dan Komposisi Peta Tematik. Semarang: CV.Indoprint.
- Liu, Huan., Ge, Yong., Zheng, Qinghua., Lin, Rongcheng., Li, Huaya. 2018. Detecting global and local topics via mining twitter data. *2018, Neurocomputing volume 273 Issue C*.
- Lo, Siaw. L., Chiong, Raymond., Comforth, David. 2017. An unsupervised multilingual approach for online social media topic identification. *2017, Expert System with Applications*.
- Medvet, Eric., Bartoli, Alberto. 2014. Brand-related events detection, classification and summarization on twitter. *2014 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*.
- Milioris, Dimitrios., Jacquet, Philippe. 2015. Topic detection and compressed classification in twitter. *2015 23rd European Signal Processing Conference (EUSIPCO)*.
- Muhammadi. 2004. Perumusan Masalah. [www.zkarnain.tripod.com](http://www.zkarnain.tripod.com). Diakses 30 desember 2018.
- Nakahara, T., Hamuro, Y. 2014. Detecting topics from twitter posts during tv program viewing. *2014 IEEE International Conference on Data Mining Workshops (ICDM)*.

Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>

- Nassar, Lobna., Ibrahim, R., Karray, F. 2016. Enhancing topic detection in twitter using the crowdsourcing process. *2016 International Conference on Collaboration Technologies and Systems (CTS)*.
- Nolasco, Diogo., Oliveira, Jonice. 2018. Subevents detection through topic modeling in social media posts. *Future Generation Computer Systems Volume 93, April 2018, Pages 290-303*.
- Nuraini, Khumaisa., najahaty, Ibtisami., Hidayati, Lina., Murfi, Hendri., Nurrohmah, Siti. 2015. Combination of singular value decomposition and K-means clustering methods for topic detection on twitter. *2015, International Conference on Advanced Computer Science and Information Systems*.
- Panem, Sandeep., Bansal, Romil., Gupta, Manish., Varma, V. 2014. Entity tracking in real time using sub-topic detection on twitter. *2014 European conf on Information Retrieval (ECIR)*.
- Paris, D. L., Bahari, M., Iahad, N., & Ismail, W. (2016). Systematic Literature Review of e-Commerce Implementations Studies. *Search Ebscohost*, 89:422-438.
- Peng, Baolin., Li, Jing., Chen, Junwen., Han, Xu., Xu, Ruifeng., Wong, Kam-Fai. 2015. Trending sentiment-topic detection on twitter. *Computational Linguistics and Intelligent Text Processing: 16th International Conference, CICLing 2015, Cairo, Egypt*.
- Petkos, Georgios., Papadopoulos, Symeon., Kompatsiaris, Yiannis. 2014. Two-level message clustering for topic detection in twitter. *2014 SNOW-DC*.
- Phuvipadawat, Swit., Murata, Tsuyoshi. 2014. Breaking news detection and tracking in twitter. *2014 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*.
- Prasad, Shiva., Melville, Prem., Banejee, Arindam., Sindhvani, V. 2014. Emerging topic detection using dictionary learning. *Proceedings of the 20th ACM Conference on Information and Knowledge Management, CIKM 2014*.
- Rahardi, Kunjara. 2009 . Bahasa Indonesia Untuk Perguruan Tinggi. Jakarta: Erlangga. Diakses 11 desember 2018.
- Ren, Yafeng., Wang, Ruimin., Ji, Donghong. 2016. A topic enhanced word embedding for twitter sentiment classification. *2016, Information Sciences*.
- Rill, Sven., Reinel, Dirk., Scheidt, Jorg., Zicari, Roberto. 2016. Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. *2016, Expert System with Applications*.
- Ristiyono. 2008. Hakikat Ilmu Pengetahuan. Jurnal.
- Saito, Shota., Tomioka, Ryota., Yamanishi, K. 2014. Early detection of persistent topic in social networks. *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*
- Saputro, F. A. 2015. Analisis Komparatif Kinerja Algoritms Latent Semantic Indexing dan K-Means Dalam Mengelompokan Dokumen Teks Pendek [skripsi]. Jakarta : Fakultas Teknik, Universitas Negeri Jakarta.
- Sharma, Pinkashia., Singh J. 2017. Systematic Literature Review on Software Effort Estimation Using Machine Learning Approaches. *International Conference on Next Generation Computing and Information Systems (ICNGCIS)*.
- Shuttleworth, Martyn. (2009). What is a Literature Review?  
<https://explorable.com/what-is-a-literature-review> . Diakses 11 Desember 2018.
- Sicilia, Rosa., Giudice, S. L., Pei, Yulong., Pechenizky, Mykola., Soda, Paolo. 2017. Health-related rumour detection on twitter. *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*.
- Soman, Saini, J., Marugappan. 2014. Detecting Malicious Tweets in Trending Topic using Clustering and Classification. *2014 International Conference on Recent Trends in Information Technology*.
- Suryadi. 2018. Apa itu Trending Topik di Twitter?  
<https://rocketmanajemen.com/apa-itu-trending-topik-di-twitter-dan-bagaimana-mencarinya/#a>. Diakses 9 Januari 2019.
- Sutrisno, Hadi. 2012. Pokok-pokok Metodologi Penelitian. Yogyakarta: Makalah.
- Umar, Husein. 2015. Riset Pemasaran dan Perilaku Konsumen. Jakarta: Gramedia
- Van, Steven., Feys, Matthias., Demeester, T., Develder, C., Schockeart, S., Dhoedt, Bart. 2014. Detection newsworthy topics in twitter. *2014 In Proceedings of the SNOW Data Challenge, Seoul, Korea*.
- Wahono, R. S. (2015). A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks. *Journal of Software Engineering*, 1:1.

Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>



- Wang, Meisong., Prakash, J. P., Solaiman, E., Lydia, Y., Zheng, L, C., Jun, Song., Ranjan, R. 2018. A multi-layered performances analysis for cloud-based topic detection and tracking in big data applications. 2018, *Future Generation Computer System*.
- Wang, Ye., Zang, Yanchun., Zhou, Bin., Jia, Yan. 2017. Semi-supervised collective matrix factorization for topic detection and document clustering. 2017 *IEEE Second International Conference on Data Science in Cyberspace (DSC)*.
- Wen, J., Li, S., Lin, Z., Hu, Y., & Huang, C. (2012). Systematic Literature Review of Machine Learning (ML) Based Software Development Effort Estimation Models. *Search Ebscohost*, 54:41-59.
- Xie, Wei., Zhu, F., Jiang, J., Lim, E., Wang K. 2014. Real-time Bursty Topic Detection from Twitter. 2014 *IEEE International Conference on Data Mining (ICDM)*.
- Xu, Kang., Qi, Gulinin., Huang, Junheng., Wu, Tiangxing., Fu, Xuefeng. 2017. Detecting bursts in sentiment-aware topics from social media. 2017, *Knowledge Based system*.
- Yin, Hongzhi., Cui, B., Lu, H., Huang, Y., Yao, J. 2014. A Unified Model for Stable and Temporal Topic Detection from Social Media Data. *International conference on data engineering*.
- Zang, Chen., Wang, hao., Cao, L., Wang, Wei., Xu, Fanjiang. 2015. A hybrid term-term relations anaysis approach for topic detection. 2015, *Knowledge Based system* 93.
- Zarrinkalam, Fattane., Fani, Hossein., Bagheri, Ebrahim., Kahani, Mohsen., Du, Weichang. 2015. Semantics-enabled user interest detection from twitter. 2015 *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*.
- Zhao, Liang., Hua, Ting., Lu, Chang., Chen, Ingray. 2016. A topic-focused trust model for twitter. 2016, *Computer Communications*.

Available at:

<http://journal.unj.ac.id/unj/index.php/pinter/article/view/23583>