



Analisis Tren Pertemuan Tatap Muka Terbatas dari Persepsi Warganet pada Twitter Menggunakan *Topic Modeling*

Akhsin Nurlayli¹⁾✉, Siti Irene Astuti Dwiningrum²⁾, Ahmad Chafid Alwi³⁾, dan Amrih Setyo Raharjo²⁾

¹Jurusan Pendidikan Teknik Elektronika dan Informatika, Fakultas Teknik, Universitas Negeri Yogyakarta, Indonesia

²Jurusan Filsafat dan Sosiologi Pendidikan, Fakultas Ilmu Pendidikan, Universitas Negeri Yogyakarta, Indonesia

³Jurusan Pendidikan Ekonomi, Fakultas Ekonomi, Universitas Negeri Yogyakarta, Indonesia

Info Artikel

Sejarah Artikel:

Diterima: November 2022

Direvisi: Januari 2023

Disetujui: Januari 2023

Keywords:

Analisis, LDA, Topic Modeling, Twitter

Abstrak

Salah satu dampak dari pandemi COVID-19 adalah pemberlakuan pembelajaran daring. Kegiatan pembelajaran daring banyak mengalami kendala, sehingga menyebabkan tidak tercapainya kompetensi pembelajaran dengan baik. Kemudian pemerintah memberlakukan program Pertemuan Tatap Muka Terbatas (PTMT), dan banyak tanggapan masyarakat terhadap program tersebut. Untuk mengetahui bagaimana persepsi masyarakat terhadap PTMT, kita dapat melakukan analisis terkait PTMT melalui media sosial. Karena media sosial merupakan salah satu media yang paling banyak digunakan oleh masyarakat selama pandemi untuk berkomunikasi, menyampaikan pendapat, mencari berita, dan lain-lain. Adapun analisis tren PTMT dari persepsi warganet yang dapat dilakukan adalah melalui *topic modeling*. Dengan mengambil data dari Twitter terkait PTMT menggunakan API (*Application Programming Interface*), selanjutnya memproses *topic modeling* dengan metode *Latent Dirichlet Allocation* (LDA), sebuah metode yang paling populer dan banyak digunakan pada penelitian *text mining*. Hasil pemodelan topik menghasilkan sepuluh klaster topik dengan nilai koheren 0,50344, dengan tiga topik yang paling banyak dibicarakan oleh warganet. Yakni tentang banyaknya tugas yang harus dikerjakan di luar jam sekolah, mereka sudah mencapai titik kebosanan menjalani sekolah dari rumah hingga menyebabkan kemalasan, dan banyak warganet yang membicarakan bahwa mereka akan melaksanakan kegiatan PTMT.

Abstract

One of the COVID-19 pandemic's impacts was online learning implementation. Online learning met many problems, and it caused many students could not accomplish their competencies. Therefore, the government enforced limited face-to-face meetings called the PTMT program, and many people responded to the program. To find out how people perceive PTMT, we can execute an analysis related to PTMT through social media. It is due to social media is one of the most widely used media by the public during the pandemic to communicate, express opinions, find news, and others. We can utilize topic modeling to analyze PTMT trends from netizens' social media posts by taking data from Twitter related to PTMT using the API, then processing topic modeling with the Latent Dirichlet Allocation (LDA) method, which is most popular and widely used in text mining research. The topic modeling results show ten topic clusters with a coherent value of 0.50344, with the three most discussed topics by netizens. Consisting of talking about the many tasks' netizens had to do exceeding school hours, they were tired of doing school from home, which caused laziness, and many netizens were talking about the upcoming PTMT.

© 2022 Universitas Negeri Semarang

✉ Alamat korespondensi:

Kompleks Fakultas Teknik, Kampus Karangmalang,
Universitas Negeri Yogyakarta, Yogyakarta, 55281
E-mail: akhsinnurlayli@uny.ac.id

ISSN 2252-6811

E-ISSN 2599-297X

PENDAHULUAN

Virus COVID-19 menyebar secara global pada akhir tahun 2019 hingga menjadi pandemi dunia pada awal tahun 2020. Mengacu pada Peraturan Menteri Kesehatan Republik Indonesia Nomor 9 Tahun 2020 Tentang Pedoman Pembatasan Sosial Berskala Besar Dalam Rangka Percepatan Penanganan Corona Virus Disease 2019 (COVID-19), hampir semua aktivitas yang mengharuskan interaksi antar manusia diberhentikan untuk menekan penyebaran virus. Kegiatan pendidikan yang dilaksanakan di berbagai sekolah dan universitas diberhentikan sementara segera setelah terdeteksinya virus COVID-19 pertama di Indonesia. Karena kegiatan-kegiatan dalam pendidikan akan memungkinkan penyebaran virus secara masif akibat interaksi manusia, maka semua kegiatan harus dilakukan secara *online* atau daring untuk mendukung program *physical distancing*. Namun ternyata dalam pembelajaran daring banyak masalah yang disebabkan oleh sistem pembelajaran jarak jauh, sebagian besar negara di dunia menghadapi masalah yang sama. Sekolah dan universitas belum cukup siap untuk sistem *online* yang meliputi materi pembelajaran, kesiapan guru dan siswa, capaian kompetensi pembelajaran, sistem pembelajaran *online*, manajemen pembelajaran, dan kasus umum lainnya dalam pembelajaran jarak jauh (Andarwulan et al., 2021; Badiozaman, 2021; Chempaka Seri Abdul Razak et al., 2021; Paliwal & Singh, 2021; Sikirit, 2020; Wang et al., 2022). Terdapat beberapa masalah pembelajaran daring yang tidak bisa terselesaikan, terutama pada capaian kompetensi pembelajaran. Sehingga pada awal tahun 2022, Indonesia melaksanakan pertemuan tatap muka terbatas pada kegiatan pembelajaran di sekolah dan universitas. Kegiatan luring mulai dilaksanakan dengan jumlah siswa/mahasiswa terbatas (Direktorat Sekolah Dasar, 2022). Siswa/mahasiswa bergantian mengikuti kegiatan pembelajaran luring.

Kenyataannya, kini perkembangan pandemi berbanding lurus dengan perkembangan media sosial. Pandemi berkembang dengan pesat penyebarannya, manusia semakin masif menggunakan media sosial dalam kesehariannya (Dewan Teknologi Informasi dan Komunikasi Nasional, 2020). Dengan demikian, media sosial menjadi salah satu sumber informasi *online* teratas yang dapat mendefinisikan masalah warga selama pandemi hampir secara *real-time* (Chempaka Seri Abdul Razak et al., 2021; Wong et al., 2021). Para pengguna media sosial yang menyampaikan kondisi mereka saat itu juga dengan mengunggah *post* berupa teks, gambar, suara, atau video pada akun mereka. Orang yang mengakses media sosial

lebih tinggi dari sebelum pandemi disebabkan adanya pemberlakuan kerja, belajar, dan kegiatan lainnya yang harus dilakukan di rumah menggunakan internet (Asghar et al., 2021; Saud et al., 2020; Wong et al., 2021). Sekarang ini, pertumbuhan media sosial untuk gaya hidup, komunitas, kesehatan, olahraga, hiburan, bisnis, dan lainnya meningkat secara signifikan (statistia.com, 2022). Hal tersebut menyebabkan data media *online* menjadi objek penelitian yang menarik di masa pandemi (Asghar et al., 2021; Lazzini et al., 2022; Lendriyono, 2021; Saud et al., 2020). Dengan adanya kebiasaan masyarakat yang lebih banyak menggunakan media sosial, kita dapat memanfaatkan data pada media sosial untuk mengetahui dan menganalisis permasalahan tertentu melalui opini publik.

Berbagai penelitian dan pengembangan telah dilakukan untuk menganalisis data media sosial, terutama pada Twitter dengan mengolah data teks di era pandemi. Seperti penelitian yang dilakukan oleh Joanne Chen Lyu, dkk (2021), penelitian tersebut mengolah data teks pada Twitter terkait vaksin COVID-19 menggunakan analisis sentimen dan *topic modeling*. Mereka menggunakan data dari Twitter yang diambil mulai tanggal 11 Maret 2020 hingga 31 Januari 2021, *tweets* yang digunakan adalah yang mengandung kata *vaccination*, *vaccinations*, *vaccine*, *vaccines*, *immunization*, *vaccinate*, dan *vaccinated*. Hasil analisis sentimen menunjukkan bahwa warganet cenderung memiliki sentimen positif dan memiliki kepercayaan yang tinggi terhadap adanya vaksin COVID-19 yang mampu mengatasi pandemi global. Hasil *topic modeling* menggunakan LDA (*Latent Dirichlet Allocation*) menunjukkan terdapat lima topik pembahasan terkait vaksin, yang meliputi opini terkait vaksin, pengetahuan tentang vaksin, vaksin sebagai isu global, administrasi dalam penyaluran vaksin, dan progres pengembangan serta otoritas terkait vaksin (Lyu et al., 2021). Chempaka Seri Abdul Razak, dkk (2021) telah melakukan penelitian untuk mendeteksi emosi pengguna Twitter terkait COVID-19 di Malaysia dengan memanfaatkan *Emotion Detection Model*, yakni mengklasifikasikan emosi pada data teks *tweet* menggunakan *Deep Learning* dengan hasil akurasi 0,9683 (Chempaka Seri Abdul Razak et al., 2021). Selanjutnya, Hyojung Kim, dkk (2022) menganalisis tren *fashion* dari persepsi konsumen pada media sosial dengan melakukan *topic modeling*, penelitian tersebut menerapkan metode LDA (*Latent Dirichlet Allocation*) untuk data mengandung kata “*genderless fashion*”. Data diambil mulai tanggal 1 Januari 2018 sampai 31 Desember 2020. Hasil penelitian yang didapatkan adalah berupa topik-

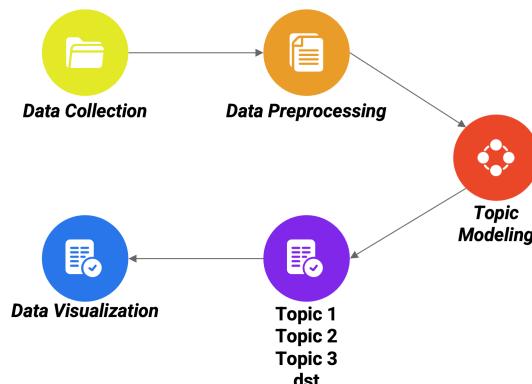
topik yang diperbincangkan oleh masyarakat terkait tren *fashion*, topik yang paling banyak dibahas adalah topik parfum dan wewangian. Topik terkait tas, *set-up suit*, dan *all-in-one skincare* adalah topik yang tidak banyak dibicarakan dan popularitasnya menurun di kalangan konsumen (Kim et al., 2022). Hasil-hasil penelitian tersebut tidak hanya bisa digunakan untuk referensi penelitian secara teknis, namun hasil penelitian yang berupa *insights* dari analisis dapat digunakan sebagai pengetahuan dan landasan dalam melakukan pengambilan keputusan untuk melakukan tindakan atau strategi lanjutan dalam *case tertentu*.

Berdasarkan analisis kondisi masyarakat dan beberapa hasil penelitian di atas, kita dapat memanfaatkan data media sosial untuk menganalisis kondisi masyarakat terkait hal tertentu yang diinginkan. Jika melihat dari kondisi di Indonesia terkini adalah tentang kebijakan pendidikan yang menjadi *trending topic* pada awal tahun 2022, yaitu Pertemuan Tatap Muka Terbatas (PTMT) untuk kegiatan pembelajaran. Kebijakan tersebut dilaksanakan atas dasar untuk mengatasi beberapa permasalahan dalam kegiatan pembelajaran daring. Meskipun PTMT dianggap sebagai solusi penyelesaian masalah pembelajaran daring untuk masyarakat, ternyata masih banyak masyarakat yang tidak setuju dengan adanya program tersebut. Untuk mengetahui bagaimana tanggapan masyarakat terhadap program PTMT, hal-hal yang menyebabkan masyarakat tidak sepakat dengan program tersebut, dan hal-hal lain terkait PTMT, peneliti memanfaatkan data Twitter (sosial media yang memiliki daya viral tinggi (Meg, 2022)) dengan melakukan *topic modeling* untuk mendapatkan data topik apa saja yang dibahas masyarakat terkait PTMT. Adapun metode yang digunakan adalah LDA, metode dalam pemodelan topik yang paling banyak digunakan dan direkomendasikan untuk melakukan pemodelan topik dalam mengolah data teks (Jelodar et al., 2018; Madzik & Falát, 2022).

METODE PENELITIAN

Topic Modeling merupakan teknik yang paling direkomendasikan untuk melakukan *text mining*, penemuan data laten, dan menemukan hubungan antara data pada data teks. Banyak penelitian terkait *topic modeling* dan diterapkan pada berbagai bidang seperti rekayasa perangkat lunak, ilmu kedokteran, linguistik, dan lain-lain. Terdapat beberapa metode dalam *topic modeling*, salah satunya adalah *Latent Dirichlet Allocation* (LDA). LDA merupakan salah satu metode yang paling populer dalam pemodelan

topik (Alghamdi & Alfalqi, 2015; Bastani et al., 2019; Garbapu, 2020; Hidayatullah et al., 2019; Jelodar et al., 2018; Kim et al., 2022; Lyu et al., 2021; Madzik & Falát, 2022; Myat Noe Win, Sri Devi Ravana, 2022; Nurlayli & Nasichuddin, 2019; Sutherland et al., 2020; Yokomoto et al., 2012; Zoghbi et al., 2016). Gambar 1 menunjukkan alur metode penelitian pemodelan topik menggunakan LDA.



Gambar 1. Alur metode penelitian *topic modeling*

Dari Gambar 1. dapat dilihat bahwa alur penelitian untuk menganalisis tren PTMT dari persepsi warganet adalah *data collection*, *data preprocessing*, *topic modeling*, pengelompokan topik, dan *data visualization*.

A. Data Collection

Dalam penelitian ini, data untuk menganalisis tren PTMT diambil dari Twitter dengan memanfaatkan API Twitter. Data diambil pada tanggal 1 Februari 2022 – 10 Maret 2022. Adapun kata kunci yang digunakan adalah “PTM, PTMT, PTMT, Pertemuan Tatap Muka, dan Pertemuan Tatap Muka Terbatas”. 1000 dataset yang diambil berupa *tweets*, *username* pengguna, waktu *tweet*, dan *author profile description*.

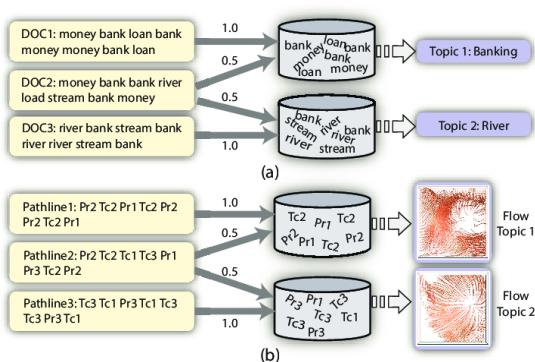
B. Data Preprocessing

Untuk menghasilkan data yang akurat, dibutuhkan adanya *data preprocessing*. Yakni: (1) Mentransformasi data *tweets* menjadi kalimat dengan lowercase; (2) Menghilangkan data *Uniform Resource Locator* (URL) pada *tweets*; (3) *Tokenization*; dan (4) Menghilangkan beberapa kata atau karakter yang tidak memiliki arti memanfaatkan *stopwords* (kata yang tidak memiliki makna) dan *regexp* (karakter yang tidak memiliki arti secara harfiah).

C. Topic Modeling

Topic Modeling adalah salah satu metode dalam analisis teks berbasis *unsupervised learning*, yakni mengelompokkan data dengan algoritma

berbasis statistik untuk menganalisis kata-kata dari teks asli, sehingga mampu mengidentifikasi pola suatu data yang besar. Dalam penelitian ini menggunakan metode *Latent Dirichlet Allocation* (LDA) untuk menganalisis tren terkait PTMT. LDA menerapkan model *generative probabilistic* untuk pengumpulan data diskrit sebagai korpus, sehingga didapatkan struktur semantik dan mendeteksi topik-topik yang ada pada koleksi data teks yang diproses dan proporsi kemunculan topik tersebut (Garbhupu, 2020; Nurlayli & Nasichuddin, 2019; Pinto & Chahed, 2015). Ilustrasi cara kerja LDA dapat dilihat pada Gambar 2.



Gambar 2. Alur proses LDA (Hong et al., 2014)

Dari Gambar 2. bisa dilihat cara kerja LDA, yaitu inisialisasi dari data teks atau dokumen, ditransformasikan menjadi dalam bentuk kata, kemudian dikelompokkan menjadi beberapa topik berdasarkan distribusi *dirichlet*. Seperti pada Gambar 2. Terdapat dokumen berisi data DOC1: *money bank loan bank money money bank loan*; DOC2: *money bank river load stream bank money*; dan DOC3: *river bank stream bank river river stream bank*. Dari ketiga dokumen tersebut dilakukan proses pembobotan LDA, sehingga ketiga dokumen tersebut dikelompokkan menjadi dua topik, yakni *Topic1: Bank* dan *Topic2: River*. Parameter yang menjadi dasar pada proses LDA dan penentuan jumlah topik yang paling baik dalam penelitian ini adalah *random state: 100, update every: 1, passes: 10, chunk size: 10, iterations: 100, chunk size: 10, alpha: symmetric, per word topics: true*, dan *coherence value: 0,50344*.

D. Data Visualization

Visualisasi data dibutuhkan untuk memudahkan dalam membaca hasil analisis *topic modeling*. Dalam penelitian ini menggunakan beberapa jenis visualisasi data sebagai berikut:

1. Word Cloud

Visualisasi *Word Cloud* digunakan untuk melihat distribusi data teks berupa kata-kata yang muncul pada *dataset* yang diolah. Pada penelitian ini berupa data kata yang muncul pada *tweets* yang diproses untuk *topic modeling*.

2. LDavis

Untuk memunculkan jenis kata dan frekuensi kata yang muncul pada pada masing-masing topik kita dapat menggunakan LDavis (*LDA visualization*), jenis visualisasi data pada LDA yang menampilkan *top ranked words*. Dengan demikian dapat dilihat kata apa saja dan berapa kali kata tersebut muncul dalam tiap topik.

3. MDS

Multidimensional Scaling (MDS) digunakan untuk menampilkan kelompok topik dalam bentuk objek, mencari hubungan antar topik, dan menunjukkan posisi antar topik berdasarkan *similarity*.

HASIL DAN PEMBAHASAN

Untuk menganalisis tren PTMT dari persepsi warganet melalui Twitter, data *tweets* akan dikelompokkan ke dalam beberapa klaster topik. Kemudian diambil tiga topik teratas yang paling banyak dibahas untuk direpresentasikan menjadi simpulan dan analisis mendalam. Dari data *tweets* yang telah diambil, Tabel 1. menunjukkan sampel kata dan frekuensi pada *dataset*.

Tabel 1. Kata dan Frekuensi Pada *Dataset*

Kata	Frekuensi
besok	180
pjj	85
sekolah	65
banget	41
bsk	41
tugas	39
kelas	36
minggu	34
jam	32
persen	31
temen	30

Berdasarkan data pada Tabel 1. Kita dapat membuat visualisasi data kata dalam bentuk *word cloud*. Gambar 3. menunjukkan *word cloud* atau kumpulan kata yang muncul pada *tweets*.



Gambar 3. Word cloud

Dari Gambar 3. dapat dilihat bahwa sepuluh kata yang paling banyak muncul adalah “besok, pjj sekolah, banget, bsk, tugas, kelas, minggu, jam, persen”. Hal ini dapat dilihat dari besarnya ukuran kata yang dimunculkan pada *word cloud*. Semakin besar ukuran kata ditampilkan, semakin besar pula frekuensi kata tersebut muncul pada data *tweets* yang kita proses.

Pemodelan topik menggunakan LDA menghasilkan lima klaster topik. Tabel 2. menunjukkan kumpulan kata pada masing-masing klaster topik.

Tabel 2. Lima Klaster Hasil Pemodelan Topik

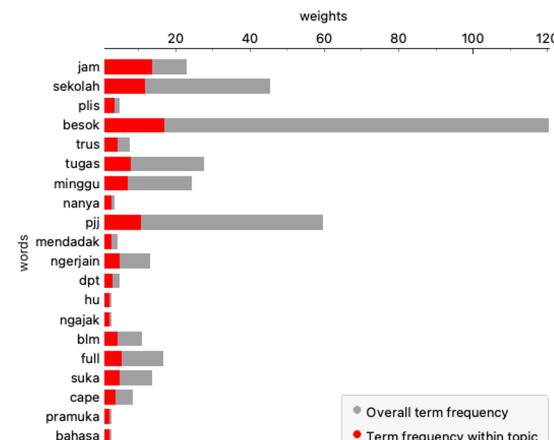
Topic	Topic Keywords
1	besok, jam, banget, semoga, hbu, tbl, senin, anak, abis, daring
2	besok, pjj, semangat, sekolah, full, ujian, pengen, langsung, level, banget
3	persen, besok, kelas, pembelajaran, tetap, pjj, muka, sekolah, suka, surabaya
4	tugas, minggu, sekolah, mingdep, besok, pulang, bsk, daring, blm, temen
5	besok, sekolah, bsk, tidur, masuk, pagi, belajar, karna, pjj, abis

Lima topik dengan distribusi kata pada masing-masing topik seperti terlihat pada Tabel 1. memiliki nilai koheren 0,49129. Untuk melakukan perbandingan dan mencari nilai koheren terbaik (untuk mendapatkan jumlah topik yang paling optimal), dilakukan iterasi beberapa kali menggunakan LDA. Pada iterasi kesekian kalinya, didapatkan sepuluh klaster topik yang menghasilkan nilai koheren tinggi (0,50344) dan memiliki distribusi *keywords* dengan pola yang baik. Distribusi *topic keywords* pada masing-masing klaster dengan jumlah sepuluh topik dapat melihat Tabel 3.

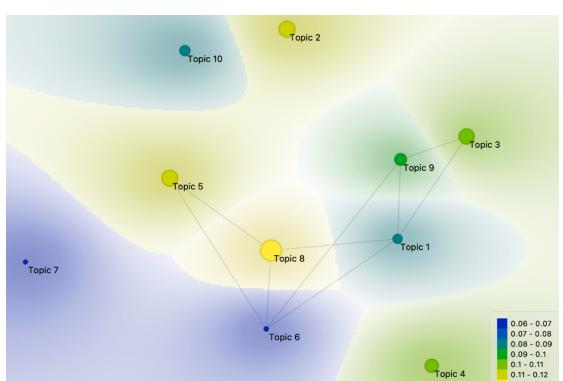
Tabel 3. Sepuluh Klaster Hasil Pemodelan Topik

Topic	Topic Keywords
1	besok, banget, kota, muka, tatap, pembelajaran, adring, batu, tbl, sekolah
2	pjj, besok, ujian, pengen, langsung, sekolah, masuk, banget, bikin, um
3	kelas, pjj, besok, terbatas, pendidikan, sekolah, temen, seneng, enakan, pembelajaran
4	mingdep, tugas, minggu, besok, sekolah, enak, buka, bsk, guys, kyk
5	besok, males, semangat, tidur, abis, sekolah, pagi, bsk, pjj, masuk
6	besok, bsk, senin, jam, sekolah, y, sesi, gws, tetep, sibuk
7	besok, pls, gempa, hbu, tu, kelas, jdi, banget, main, liat
8	besok, jam, sekolah, pjj, tugas, minggu, temen, banget, full, kelas
9	besok, seru, tatap, muka, ko, pts, temen, kelas, persen, sekolah
10	besok, persen, level, pjj, ppkm, surabaya, takut, banget, diterapkan, kota

Setelah dihasilkan pemodelan topik ke dalam sepuluh klaster seperti yang terlihat pada Tabel 3, kita dapat melihat distribusi frekuensi *keyword* pada masing-masing topik menggunakan LDAvis. Misalnya untuk melihat berapa kali kata “besok” muncul pada *Topic 8* (lihat Gambar 4.).

Gambar 4. Frekuensi setiap kata pada *topic 4*

Pada Gambar 4. dapat dilihat bahwa kata “besok” muncul hampir dua puluh kali pada *Topic 8* dari jumlah kemunculannya pada seluruh dataset *tweets*. Selain melihat jumlah frekuensi kata muncul pada masing-masing topik, kita dapat melihat topik yang paling banyak dibahas dan hubungan antar topik. Dengan memvisualisasikan melalui MDS seperti pada Gambar 5.



Gambar 5. Visualisasi topik menggunakan MDS

Berdasarkan Gambar 5., dapat dilihat bahwa ukuran objek topik yang paling besar adalah *Topic 8*, *Topic 5*, dan *Topic 2*. Hal ini menunjukkan bahwa ketiga klaster topik tersebut merupakan topik yang paling banyak dibahas oleh warganet. Adapun wawasan lain yang didapatkan dari visualisasi tersebut adalah adanya hubungan antar topik, artinya topik-topik tersebut memiliki similaritas atau kemiripan. Gambar 4. di atas memperlihatkan bahwa ada hubungan antara *Topic 1*, *Topic 3*, *Topic 5*, *Topic 6*, *Topic 8*, dan *Topic 9*. Adapun rincian data *tweets* pada masing-masing topik dapat dilihat pada Gambar 5., misalnya jika kita ingin melihat *tweets* pada klaster *Topic 8*.

Tabel 3. Sampel *Tweets Detail* Pada *Topic 8*

No	<i>Tweet</i>
1	@tanyakanrl WEH AKU PERNAH.ayah ada urusan pago pagi, hari itu aku sekolah ptm.trus sampe jam 7 ga pulang pulang aku mulai baris di...
2	di gcr Cuma ngasih materi, dan gadisuruh catet.tpi guw ttp nyatet.gue nih bisa dibilang pj paling santai diantara pj lainnya.soalnya bu pkn.....
3	Jadwal PTM pun nanya gue,kadang dahal uda gue jawab "Lho di edaran ka nuda tertera BU" masih aja ngajak ngobrol gajelas....
4	@schfess Jadi sekolah aku diadain ptm 50% lagi,dan lagi karena walikelas aku lambat ngasih info, aku dpt lagi info dari kelas sebelah dan aku.....
5	Hari ketiga di minggu keempat,tskut besok udah mulai PTM dan ad akelas pagi jam 7...takut tbtb disuruh ngajar...

Rincian data *tweets* juga dapat dilihat untuk klaster topik yang lainnya. Tabel 3. adalah sampel dari *tweets* yang masuk dalam klaster *Topic 8*. Selanjutnya, kita perlu menginterpretasikan

topik melalui *topic keywords* dan data *tweets* pada masing-masing klaster topik. Melihat pada Tabel 2. dan data *tweets* pada masing-masing klaster, kita bisa menarik kesimpulan bahwa: (1) *Topic 8* berisi kata "besok, jam, sekolah, pjj, tugas, minggu, temen, banget, full, kelas". Dari data *tweets* dan kata yang muncul pada klaster tersebut dapat diinterpretasikan bahwa topik tersebut berbicara tentang banyak tugas yang harus dikerjakan di luar jam sekolah; (2) *Topic 5* berisi kata "besok, males, semangat, tidur, abis, sekolah, pagi, bsk, pjj, masuk". Dari data *tweets* dan kata yang muncul pada klaster tersebut dapat diinterpretasikan bahwa warganet sudah mencapai titik kebosanan menjalani sekolah dari rumah, sehingga menimbulkan kemalasan; dan (3) *Topic 2* berisi kata "pjj, besok, ujian, pengen, langsung, sekolah, masuk, banget, bikin, um". Dari data *tweets* dan kata yang muncul pada klaster tersebut dapat diinterpretasikan terkait akan dilaksanakannya PTMT.

A. Persepsi Warganet Terkait Performa Pembelajaran Jarak Jauh (PJJ)

Berbagai data visual yang dihasilkan dari *Topic Modeling* mengindikasikan beberapa poin reflektif. Poin-poin tersebut setidaknya menjelaskan dua hal: 1) Bagaimana persepsi warganet mengenai performa PJJ; dan 2) *Lesson learned* yang dapat diambil dari PJJ untuk mempersiapkan PTMT. Dua hal tersebut secara implisit tercermin dari data-data visual yang telah terhimpun. Seperti visualisasi *Word Cloud* pada Gambar 2. yang menjelaskan indikasi antusiasme menjelang PTMT. Frekuensi setiap kata pada Gambar 3. memberikan gambaran tentang respon warganet mengenai perbandingan PJJ dan PTMT. Serta Sampel *tweets detail* yang mencerminkan sentimen warganet.

Berbagai visualisasi data pada hasil *topic modeling* dan memberikan sudut-sudut gambaran baru mengenai bagaimana persepsi masyarakat implementasi PJJ. PJJ masa COVID-19 terjadi dengan latar belakang yang tidak biasa. Ia dipaksakan teraplikasi karena fenomena *black swan* dari pandemi. COVID-19 diibaratkan sebagai angsa hitam (*black swan*) di tengah kelumrahan angsa putih di masyarakat. Kajian Antipova (2020) menyatakan bahwa, "Coronavirus Pandemic recognized as Black Swan event according to considered figures and facts from recent references". COVID-19 adalah fenomena baru di mana peradaban belum menemukan penawarnya. Sementara itu, pendidikan harus terus berjalan di tengah ketidaktahuan. Oleh karena itu, Pendidikan dijalankan dengan terpaksa dan serba tanpa persiapan (Antipova,

2020). Sehingga, bagaimana PJJ terlaksana selama pandemi dapat dilihat dari hasil *topic modeling* dan visualisasi data. Hasil penelitian mampu menggambarkan respon warganet terhadap pelaksanaan PJJ. Hasil *text mining* menemukan pola bahwa pada konteks PJJ, pengajar masih gagap dalam melaksanakan PJJ. Pengajar belum mampu mendesain pembelajaran berbasis digital yang baik. Sehingga warganet merasa PJJ tidak efektif karena terjadi penumpukan tugas dan materi, tanpa dibarengi dengan pemahaman. Kegagalan PJJ ini diafirmasi dengan salah satu hasil *text mining* menyatakan bahwa, " ... di GCR (Google classroom) Cuma ngasih materi ...". *Tweet* tersebut mengindikasikan bahwa pembelajaran menjadi tidak interaktif karena pengajar terbatas dalam penggunaan media belajar.

Fenomena kegagalan belajar berbasis PPJ diafirmasi oleh kajian Basar (2021) mengenai praktik PJJ di SMP. Kajian Basar menerangkan bahwa proses PJJ terkendala baik dari pendidik maupun siswa. Guru terbatas dalam mengoperasikan media belajar, sementara siswa kehilangan motivasi belajar. Berdasar problematika PJJ hasil *text mining* dan penelitian terkait, perlu adanya peningkatan skill pengajar dalam menerangkan materi agar pembelajaran menjadi interaktif. Sementara itu, perlu juga pembelajaran yang memang didesain secara daring agar motivasi anak dapat terjaga untuk mengikuti pembelajaran via *online* (Basar, 2021).

B. PTMT dan Pembelajaran Pasca Pandemi

PTMT ialah transisi menuju pembelajaran pasca pandemi. Pembelajaran setelah pandemi menjadi endemik dan terkontrol. Saat sekolah dan masyarakat dapat melangsungkan pembelajaran secara luring penuh, sekaligus tahu bagaimana cara menjalankan pembelajaran secara daring. Lalu, bagaimana seharusnya pembelajaran pasca pandemi terlaksana? Hasil *text mining* menandakan bahwa warganet antusias sekaligus bingung dalam menghadapi PTMT. Reaksi ini tercermin dalam percakapan-percakapan di platform twitter. Kata "Besok" muncul sebagai kata yang sering dituliskan. Kata "Besok" muncul dalam pencarian topik PTMT menandakan warganet memberi perhatian pada fenomena tersebut. Apabila dirincikan lagi melalui sampel *tweet*, warganet menyampaikan setimen positif maupun negatif. Sentimen positif seperti, "... besok jadwal ptm, gue mau masuk...", atau "...mending ptm aja semua..." *tweet-tweet* tersebut menandakan reaksi memilih PTM dibandingkan PJJ.

PTMT didesain untuk merespon masalah pembelajaran daring. Akan tetapi, PTMT bukanlah pembelajaran yang sepenuhnya luring. Sehingga desain pembelajaran campuran perlu dirumuskan secara baik. Pembelajaran gabungan ini disebut sebagai *blended learning*. Pembelajaran ini ialah gabungan model belajar format *online* dan *offline* dengan komposisi yang beragam (Abdullah, 2018). *Blended learning* dapat menjadi alternatif pembelajaran pasca pandemi. *Blended learning* menawarkan pembelajaran variatif dan interaktif karena mampu memberikan pemahaman serta efektifitas dengan bantuan teknologi. Namun, untuk mewujudkan keadaan tersebut perlu adanya desain pembelajaran *blended*. Pendekatan Desain pembelajaran ini akan memberikan pemahaman melalui komunikasi interaktif karena mengoptimalkan berbagai fitur dalam *platform* daring. .

SIMPULAN

Penelitian ini telah menerapkan *topic modeling* dengan menggunakan metode LDA (*Latent Dirichlet Allocation*) untuk menganalisis tren Pertemuan Tatap Muka Terbatas (PTMT) dari persepsi warganet. PTMT merupakan program pemerintah untuk mengatasi beberapa masalah pada pembelajaran daring. LDA merupakan salah satu metode yang paling populer dan mampu melakukan pemodelan topik dengan baik yang terbukti melalui beberapa penelitian *topic modeling*. Hasil penelitian yang telah dilakukan menunjukkan bahwa sepuluh klaster merupakan pengelompokan topik terbaik dan memiliki nilai koheren tinggi (0,50344). Dari sepuluh klaster topik tersebut, terdapat tiga topik teratas yang paling banyak dibahas, yakni tentang banyaknya tugas yang harus dikerjakan warganet di luar jam sekolah, warganet sudah mencapai titik kebosanan menjalani sekolah dari rumah hingga menimbulkan kemalasan, dan banyak warganet yang membicarakan bahwa mereka akan melaksanakan kegiatan PTMT dalam waktu dekat. Berbagai data tersebut mencerminkan kendala dalam proses pembelajaran daring. Kegagalan pengajar mendesain pembelajaran daring menyebabkan proses belajar berlangsung tidak efektif. Pembelajaran daring masih dilakukan sebatas meng-*online-kan* materi yang selama ini disampaikan secara *offline*. Hal tersebut tidak kompatibel dengan model belajar daring yang memerlukan interaksi secara kontinu untuk menjaga fokus peserta didik. Adapun penelitian lanjutan untuk analisis tren berdasarkan persepsi masyarakat melalui twitter adalah dengan melakukan optimalisasi *data preprocessing* untuk

mengatasi *tweets* yang kurang relevan, misalnya *tweets* oleh *buzzer* atau penjual *online* yang menggunakan *keyword(s)* tidak relevan untuk meningkatkan *engagement* akun.

DAFTAR PUSTAKA

- Abdullah, W. (2018). *Model Blended Learning dalam Meningkatkan Efektifitas Pembelajaran*. 7.
- Alghamdi, R., & Alfalqi, K. (2015). A Survey of Topic Modeling in Text Mining. *International Journal of Advanced Computer Science and Applications*, 6(1), 147–153. <https://doi.org/10.14569/ijacsa.2015.060121>
- Andarwulan, T., Al Fajri, T. A., & Damayanti, G. (2021). Elementary teachers' readiness toward the online learning policy in the new normal era during Covid-19. *International Journal of Instruction*, 14(3), 771–786. <https://doi.org/10.29333/iji.2021.14345a>
- Antipova, T. (2020). *Integrated Science in Digital Age* 2020. https://doi.org/doi.org/10.1007/978-3-030-49264-9_32
- Asghar, M. Z., Iqbal, A., Seitamaa-Hakkarainen, P., & Barbera, E. (2021). Breaching learners' social distancing through social media during the covid-19 pandemic. *International Journal of Environmental Research and Public Health*, 18(21). <https://doi.org/10.3390/ijerph182111012>
- Badiozaman, I. F. A. (2021). Exploring online readiness in the context of the COVID 19 pandemic. *Teaching in Higher Education*, 0(0), 1–19. <https://doi.org/10.1080/13562517.2021.1943654>
- Basar, A. M. (2021). *Problematika Pembelajaran Jarak Jauh Pada Masa Pandemi Covid-19 (Studi Kasus di SMPIT Nurul Fajri – Cikarang Barat – Bekasi)*. 2(1), 208–218.
- Bastani, K., Namavari, H., & Shaffer, J. (2019). Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints. *Expert Systems with Applications*, 127, 256–271. <https://doi.org/10.1016/j.eswa.2019.03.001>
- Chempaka Seri Abdul Razak, Hamid, S. H. A., Meon, H., Subramaniam, H. A., & Anuar, N. B. (2021). TWO-STEP MODEL FOR EMOTION DETECTION ON TWITTER USERS: A COVID-19 CASE STUDY IN MALAYSIA. *Malaysian Journal of Computer Science*, 34(4), 374–388.
- Dewan Teknologi Informasi dan Komunikasi Nasional. (2020). *Akses Digital Meningkat Selama Pandemi*. <http://www.wantiknas.go.id/id/berita/akses-digital-meningkat-selama-pandemi>
- Direktorat Sekolah Dasar. (2022). *Semua Sekolah Wajib Melaksanakan PTM Terbatas pada 2022*.
- Garbhanu, V. K. (2020). A comparative analysis of Latent Semantic analysis and Latent Dirichlet allocation topic modeling methods using Bible data. *Indian Journal of Science and Technology*, 13(44), 4474–4482. <https://doi.org/10.17485/ijst/v13i44.1479>
- Hidayatullah, A. F., Aditya, S. K., Karimah, & Gardini, S. T. (2019). Topic modeling of weather and climate condition on twitter using latent dirichlet allocation (LDA). *IOP Conference Series: Materials Science and Engineering*, 482(1). <https://doi.org/10.1088/1757-899X/482/1/012033>
- Hong, F., Hong, F., Lai, C., Guo, H., Member, S., Shen, E., Yuan, X., & Li, S. (2014). *FLDA: Latent Dirichlet Allocation Based Unsteady Flow Analysis* *FLDA: Latent Dirichlet Allocation Based Unsteady Flow Analysis*. December. <https://doi.org/10.1109/TVCG.2014.2346416>
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2018). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78, 183–198.
- Kim, H., Cho, I., & Park, M. (2022). Analyzing genderless fashion trends of consumers' perceptions on social media: using unstructured big data analysis through

- Latent Dirichlet Allocation-based topic modeling. *Fashion and Textiles*, 9(1). <https://doi.org/10.1186/s40691-021-00281-6>
- Lazzini, A., Lazzini, S., Balluchi, F., & Mazza, M. (2022). Emotions, moods and hyperreality: social media and the stock market during the first phase of COVID-19 pandemic. *Accounting, Auditing and Accountability Journal*, 35(1), 199–215. <https://doi.org/10.1108/AAAJ-08-2020-4786>
- Lendriyono, F. (2021). Public's Perception on Social Media towards New Normal during Covid-19 Pandemic in Indonesia: Content Analysis on Religious Social Media Accounts. *IOP Conference Series: Earth and Environmental Science*, 717(1). <https://doi.org/10.1088/1755-1315/717/1/012039>
- Lyu, J. C., Han, E. Le, & Luli, G. K. (2021). Covid-19 vaccine-related discussion on twitter: Topic modeling and sentiment analysis. *Journal of Medical Internet Research*, 23(6). <https://doi.org/10.2196/24435>
- Madzik, P., & Falát, L. (2022). State-of-the-art on analytic hierarchy process in the last 40 years: Literature review based on Latent Dirichlet Allocation topic modelling. In *Plos One* (Vol. 17, Issue 5). <https://doi.org/10.1371/journal.pone.0268777>
- Meg. (2022). *Social media listening guide*. <https://www.talkwalker.com/blog/social-media-listening-guide>
- Myat Noe Win, Sri Devi Ravana, and L. S. (2022). SENTIMENT ATTRIBUTION ANALYSIS WITH HIERARCHICAL CLASSIFICATION AND AUTOMATIC ASPECT CATEGORIZATION ON ONLINE USER REVIEWS. *Malaysian Journal of Computer Science*, 35(2), 89–110.
- Nurlayli, A., & Nasichuddin, M. A. (2019). *Topic Modeling Penelitian Dosen JPTEI UNY pada Google Scholar Menggunakan Latent Dirichlet Allocation*. 4(November), 154–161. <https://doi.org/10.21831/elinvo.v4i2>
- Paliwal, M., & Singh, A. (2021). Teacher readiness for online teaching-learning during COVID-19 outbreak: a study of Indian institutions of higher education. *Interactive Technology and Smart Education*, 18(3), 403–421. <https://doi.org/10.1108/ITSE-07-2020-0118>
- Pinto, J. C. L., & Chahed, T. (2015). Modeling multi-topic information diffusion in social networks using latent dirichlet allocation and hawkes processes. *Proceedings - 10th International Conference on Signal-Image Technology and Internet-Based Systems, SITIS 2014*, 339–346. <https://doi.org/10.1109/SITIS.2014.24>
- Saud, M., Mashud, M., & Ida, R. (2020). Usage of social media during the pandemic: Seeking support and awareness about COVID-19 through social media platforms. *Journal of Public Affairs*, 20(4). <https://doi.org/10.1002/pa.2417>
- Sikirit, D. (2020). *The Challenges of Home Learning during the COVID-19 Pandemic | UNICEF Indonesia*. <https://www.unicef.org/indonesia/education-and-adolescents/coronavirus/stories/learning-home-during-covid-19-pandemic>
- statista.com. (2022). *Number of social media users worldwide from 2017 to 2027*. <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>
- Sutherland, I., Sim, Y., Lee, S. K., Byun, J., & Kiatkawsin, K. (2020). Topic modeling of online accommodation reviews via latent dirichlet allocation. *Sustainability (Switzerland)*, 12(5), 1–15. <https://doi.org/10.3390/su12051821>
- Wang, Y., Xia, M., Guo, W., Xu, F., & Zhao, Y. (2022). Academic performance under COVID-19: The role of online learning readiness and emotional competence. *Current Psychology*, 0123456789. <https://doi.org/10.1007/s12144-022-02699-7>
- Wong, A., Ho, S., Olusanya, O., Antonini, M. V., & Lyness, D. (2021). The use of social media and online communications in times of

pandemic COVID-19. *Journal of the Intensive Care Society*, 22(3), 255–260.
<https://doi.org/10.1177/1751143720966280>

Yokomoto, D., Makita, K., Suzuki, H., Koike, D., Utsuro, T., Kawada, Y., & Fukuhara, T. (2012). LDA-based topic modeling in labeling blog posts with wikipedia entries. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7234 LNCS, 114–124.
https://doi.org/10.1007/978-3-642-29426-6_15

Zoghbi, S., Vulić, I., & Moens, M. F. (2016). Latent Dirichlet allocation for linking user-generated content and e-commerce data. *Information Sciences*, 367–368, 573–599.
<https://doi.org/10.1016/j.ins.2016.05.047>