

Klasterisasi Topik Konten Channel Youtube Gaming Indonesia Menggunakan Latent Dirichlet Allocation

Arief Rahman¹, Rekyan Bayu Waskitho², Muhammad Fajrul Alam Ulin Nuha³, Nur Aini Rakhmawati⁴

^{1,2,3,4}Program Studi S1 Sistem Informasi, Institut Teknologi Sepuluh Nopember

³fajrul.19052@mhs.its.ac.id

Abstrak— Youtube adalah platform untuk saling berbagi video terbesar di internet. Semakin platform ini berkembang, semakin banyak konten yang tersedia di dalamnya, yang dikarenakan semakin beragam genre videonya. Salah satu genre video yang sedang naik daun adalah konten *gaming*, yang mana topik tersebut adalah objek pada penelitian ini. Penelitian ini dilakukan dengan menggunakan metode *Latent Dirichlet Allocation* (LDA) untuk memetakan topik-topik dari genre *gaming* ini. Data didapatkan dari 10 kanal *gaming* dengan *subscribers* terbanyak di Indonesia, yang diekstrak dengan melakukan *text mining*. Total data yang didapatkan adalah sebanyak 12.757 video. Sekian banyak video ini dipetakan menjadi 5 topik yang paling dominan. Adapun jumlah topik yang diambil didasarkan pada perhitungan *perplexity*, dan keterkaitan kata dalam topik dihitung menggunakan *coherence*. Topik tersebut antara lain bocoran *update* dan *season* baru, *review hero* dan *skin* baru, *live* bermain *game* di saat ada *event* yang menarik, membahas *bug* yang terjadi pada *patch* baru rilis sehingga mempengaruhi *gameplay*, dan konten memborong *skin* dalam *game*.

Kata Kunci— *Latent Dirichlet Allocation*, *Gaming*, analisis topik, topik dominan, konten *youtube*

I. PENDAHULUAN

Youtube adalah situs yang menyediakan layanan video di internet. Pertama kali didirikan pada tahun 2005 oleh tiga mantan pegawai *PayPal* [7]. Situs ini memberi kebebasan bagi siapapun untuk mengunggah videonya agar bisa ditonton oleh masyarakat di seluruh dunia.

Saat ini *youtube* menyanggah gelar penyedia layanan video terbesar di dunia. Di Indonesia, *youtube* adalah media sosial yang paling populer, dengan 88% penduduk Indonesia merupakan pengguna *youtube* [10]. Berdasarkan penelitian yang dilakukan oleh Golnari dkk., cengkraman *youtube* di ajang dominasi media sosial akan terus menguat. Perkembangan *youtube* yang pesat ini tidak lepas dari kontribusi para konten kreator [9]. Seiring berkembangnya platform *youtube*, konten yang dituangkan juga semakin lengkap dan bervariasi. Salah satu jenis konten yang saat ini sedang naik daun adalah konten *gaming*.

Game online kini telah menjadi tren yang banyak diminati tidak hanya dapat bermain sendiri (*single*), namun juga memungkinkan bermain bersama pengguna lain dari lokasi yang berbeda (*multiplayer*). Saat ini pemain *game* telah dapat bermain bersama pengguna lain dari lokasi yang berbeda, bahkan antar kota hingga dapat antar negara. Berbagai *game* yang membutuhkan jaringan internet (*online games*) ini ternyata memiliki banyak peminat dan mempunyai penggemar yang begitu banyak. Sering kali kita menjumpai di tempat-

tempat persewaan *game* (*game centre*) yang tidak pernah sepi pengunjung, Bahkan pada waktu tertentu terdapat banyak orang yang mengantri untuk dapat bermain *game online* [1].

Game pada dasarnya merupakan sebuah interaktif, aktivitas yang berpusat dalam sebuah pencapaian dan terdapat pelaku aktif. Dalam *game* terdapat aturan yang bertujuan untuk membatasi perilaku pemain sekaligus menentukan permainan [4]. Karena melihat *game* sebagai ladang yang bagus untuk berkarya maka juga berdampak pada *youtuber gaming* yang semakin banyak bermunculan. Konten yang dibuat oleh para *youtuber gaming* cukup diminati oleh banyak penonton. Dengan semakin banyaknya orang yang tertarik dengan konten *game*, memahami pola penonton dan juga konten yang sering ditonton atau disukai dapat membantu para konten kreator untuk membuat konten yang berkualitas dan lebih menarik banyak penonton. Guna mendukung para kreator konten yang akan terjun di dunia *youtube* untuk berkontribusi dalam konten *game*, maka penting untuk mengetahui topik konten yang diminati oleh penonton.

Youtube merupakan salah satu tempat yang cocok untuk pemain *game*. Melalui *youtube* mereka dapat memperoleh sekaligus berbagi ilmu tentang *game* populer saat ini. Dengan adanya *youtuber gaming*, dapat memberikan ilmu sekaligus hiburan untuk penonton *youtube* yang tertarik dengan *game* tersebut. Terdapat banyak kreator *game* di Indonesia. Saat ini, di Indonesia cukup banyak konten kreator *game* mulai dari anak-anak, remaja, hingga orang dewasa [14].

Menurut data di website social blade, salah satu *youtuber gaming* dengan *subscribers* terbanyak di Indonesia yaitu Jess No Limit, merupakan peringkat 3 *Youtuber* terbesar di Indonesia dan peringkat 13 *youtuber* di dunia pada kategori *gaming*. Maka dari itu peneliti mengamati bahwa konten *gaming* cukup diminati oleh penonton *youtube* di Indonesia bahkan di dunia.

Penelitian ini akan memanfaatkan metode *Latent Dirichlet Allocation* (LDA) untuk menganalisis konten *channel youtube gaming indonesia*. LDA akan memodelkan judul-judul konten yang diteliti sebagai campuran acak atas topik tidak terlihat (Suhartono, 2018). Dalam kalimat yang lebih sederhana: model LDA berusaha untuk merepresentasikan setiap konten menjadi beberapa topik, dan setiap topik terdiri dari distribusi kata-kata. Dengan model ini nanti dapat dihasilkan beberapa analisa menarik, diantaranya: mana topik yang paling digemari masyarakat, mana topik yang paling *mainstream*, mana topik yang digemari masyarakat namun kurang dieksploitasi.

Diharapkan dari penelitian ini dapat memberikan manfaat bagi konten kreator agar mengetahui topik dominan yang sering

diunggah oleh *youtuber gaming*, sehingga dapat menjadi referensi bagi konten kreator yang baru akan mulai berkarya. Selain itu praktisi akademik dapat menjadikan karya tulis ilmiah ini sebagai bahan rujukan untuk penelitian yang berkaitan.

Sejumlah penelitian telah menunjukkan penggunaan LDA sebagai cara untuk memahami topik yang sedang tren di media sosial. Sebuah studi tentang pemodelan topik *K-Pop* memberikan wawasan yang berguna bagi pembuat konten untuk menyelaraskan konten mereka dengan sentimen populer [17]. Studi lain juga menunjukkan bahwa LDA membantu mengidentifikasi kesamaan dan korelasi antara setiap postingan di instagram yang dielaborasi dengan posting sampel untuk memberikan wawasan dan harapan publik terhadap hukum [11]. LDA juga membantu peneliti untuk mengidentifikasi topik penelitian yang terkait dengan perilaku siswa dalam penggunaan media pembelajaran secara daring [14]. Dengan tetap menggunakan metode yang sama yakni LDA dan menggunakan data utama yang diambil adalah judul dan deskripsi video untuk menentukan topik, penelitian ini diharapkan bisa memberikan gambaran terutama bagi para kreator baru yang ingin masuk ke dunia *youtube* dengan konten *gaming*.

II. TINJAUAN PUSTAKA

A. Game

Game dapat berarti sebagai sebuah arena keputusan dan aksi pemainnya. Ada tujuan dan misi yang harus dilakukan untuk dapat mencapai target tertentu. Kelincahan intelektual, pada tingkat tertentu merupakan ukuran sejauh mana *game* itu menarik untuk dimainkan secara maksimal [6]. *Game* atau permainan adalah sesuatu yang dapat dimainkan dengan aturan tertentu sehingga ada yang menang dan ada yang kalah, biasanya dalam konteks tidak serius dengan tujuan *refreshing* atau hiburan [15].

B. Youtuber Gaming

Youtube merupakan situs web yang memungkinkan penggunaannya untuk mengunggah dan menonton video. Teknologi yang digunakan *youtube* adalah *Adobe Flash Video* dan *HTML5* yang digunakan untuk menampilkan video yang dibuat para kreator. *youtuber* merupakan sebutan untuk orang yang membuat sebuah video kemudian mengunggahnya di *youtube* sehingga dapat ditonton oleh banyak orang [12].

Ada banyak konten video yang dibuat di *youtube* seperti klip film, klip televisi, video musik, vlog, film pendek, video edukasi, sampai konten *gaming*. Pembuat konten *youtube* yang membuat video tentang *game* sering disebut sebagai *youtuber gaming* [13]. Seperti kreator *youtube* pada umumnya, *youtuber gaming* biasa membuat video seputar *game* dan biasanya melakukan *live streaming* di *youtube*.

C. Text Mining

Text mining adalah salah satu proses *data mining* berupa teks atau tulisan yang mana data yang diambil berasal dari dokumen. Tujuan dari dilakukannya *text mining* ini adalah menemukan kata-kata yang dapat mewakili sebagian besar isi dokumen sehingga data yang didapatkan dapat dianalisa hubungannya dengan dokumen yang lain. Sumber data yang digunakan menggunakan proses *text mining* ini merupakan dokumen yang formatnya tidak terstruktur, bentuk yang tidak jelas, dan sulit diuraikan dengan algoritma. Adapun dokumen yang digunakan pada *text mining* ini dapat berupa dokumen statis maupun dokumen dinamis [8].

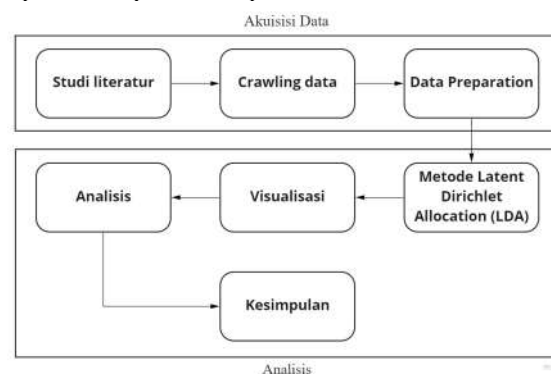
D. Metode Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) adalah suatu model *machine learning* yang digunakan untuk mengelompokkan topik-topik secara otomatis. LDA menerima input berupa judul-judul dokumen, kemudian menemukan pola dari kemunculan kata-kata yang terdapat dalam dokumen yang dimasukkan. Dari pola-pola tersebut, akan dibuatkan sejumlah topik secara otomatis. LDA adalah model yang menggunakan teknik *unsupervised learning*. Maka dari itu, semakin banyak dokumen yang dimasukkan, maka semakin akurat model yang dihasilkan oleh algoritma ini.

III. METODOLOGI

A. Alur Penelitian

Penelitian ini akan melakukan dua tahapan penelitian untuk dapat mencapai hasil akhir sesuai dengan tujuan penelitian yakni akuisisi data dan analisis. Pada tahap akuisisi data terdapat tiga proses yang akan dilakukan dalam penelitian ini dimulai dari studi literatur, data *crawling*, data *preparation*. Pada tahap kedua yaitu analisis meliputi proses LDA, visualisasi, analisis, dan kesimpulan. Metode yang kami gunakan dalam penelitian ini yaitu metode *Latent Dirichlet Allocation* (LDA). Data yang digunakan akan diambil dari judul video *youtube* berkategori *gaming* yang dibuat oleh beberapa *youtuber gaming* yang kami pilih. Gambar 1 berikut merupakan alur proses dari penelitian kami.



Gbr. 1 Alur Penelitian

B. Studi Literatur

Secara Umum Studi Literatur adalah cara untuk menyelesaikan persoalan dengan menelusuri sumber-sumber tulisan yang pernah dibuat sebelumnya. Dalam sebuah penelitian yang hendak dijalankan, tentu saja seorang peneliti harus memiliki wawasan yang luas terkait objek yang akan diteliti. Oleh karena itu, langkah awal dalam penelitian ini adalah melakukan studi literatur dan juga penggalian informasi terkait dengan metode dan cara kerja dari jurnal terdahulu yang berkaitan dengan konten penulisan ini sebagai referensi dalam penelitian. Studi literatur dilakukan dengan merujuk pada jurnal, artikel ilmiah, skripsi dan website resmi terpercaya yang dapat mendukung berjalannya penulisan ilmiah kali ini. Beberapa informasi yang didapatkan meliputi penjelasan mengenai *youtube*, *game*, pengertian LDA hingga metode penelitian.

C. Crawling Data

Crawler web adalah metode penting untuk mengumpulkan data dan mengikuti perkembangan Internet yang sangat pesat. *Crawler web* juga bisa disebut sebagai masalah pencarian grafik karena web dianggap sebagai grafik yang besar yang direpresentasikan sebagai kumpulan node. *Crawler web* dapat digunakan di berbagai bidang, yaitu salah satu yang paling menonjol adalah mengindeks satu set halaman yang besar dan memungkinkan orang lain untuk mencari indeks ini.

URL *youtube* akan menjadi acuan untuk mencari semua *hyperlink* yang kemudian akan dilakukan indexing untuk mencari data dalam dokumen pada setiap link ada (Nikmah, 2021). Data yang akan dicari merupakan data yang berhubungan dengan topik *game* di beberapa *channel youtube gaming*. *Crawling* data kemudian akan menghimpun terlebih dahulu 10 akun *youtuber gaming* dengan *subscriber* terbanyak yang nantinya dijadikan target *crawling*. Nantinya akan dilakukan *data mining* untuk mendapatkan data berupa judul, deskripsi, jumlah *subscribers*, dan jumlah video yang akan dihimpun untuk proses penelitian berikutnya.

D. Data Preparation

Data yang diperoleh perlu dibersihkan agar model yang dihasilkan representatif terhadap dunia nyata. Proses *Data Preparation* mencakup menstandarisasi data, menghapus kesalahan seperti duplikasi, kesalahan eja, atau data kosong. Data yang sejenis dimasukkan dalam suatu indeks yang membedakan dokumen satu dengan lainnya [16]. Selain itu, teks akan dipisahkan menjadi unit-unit kecil yang disebut token. Proses ini disebut dengan tokenisasi.

E. Penerapan LDA

Setelah pembersihan dilakukan, barulah diterapkan metode LDA. Tahap pertama adalah menentukan parameter. Kemudian akan dilakukan semi *random distribution*, dan dilanjutkan dengan iterasi yang menentukan distribusi dokumen. Adapun jumlah topik dioptimal dengan perhitungan *perplexity*. Rendanya nilai *perplexity* mengindikasikan model LDA yang baik [16].

F. Visualisasi

Data hasil dari penerapan LDA akan menjadi data mentah yang kemudian diolah menjadi sebuah visualisasi dengan memanfaatkan grafik dan *bar chart* untuk mempermudah proses analisis data berikutnya. Tujuan utama visualisasi data adalah untuk mengkomunikasikan informasi secara jelas dan efisien melalui grafik statistik, plot, dan grafik informasi.

G. Analisis

Pada metode LDA akan menghasilkan data mentah beserta visualisasi data sebelumnya berupa grafik dan *barchart* yang akan menjadi input dalam proses analisis ini. Proses analisis ini akan menggunakan metode *topic modelling*. Metode ini akan mengelompokkan topik tertentu dari *raw data* berdasarkan kesamaan atau kata kunci tertentu. Sedangkan, grafik yang telah ditemukan pada tahap analisis akan terus dievaluasi untuk menemukan topik yang dominan terhadap suatu konten gaming dan juga konten yang dominan dalam setiap *channel* yang telah dipilih. [16]

H. Kesimpulan

Data yang telah melalui proses analisis akan menghasilkan kesimpulan dari 10 *channel youtube* dengan kategori gaming yang mana telah digunakan sebagai subjek penelitian. Dari analisis, akan dapat ditarik kesimpulan berupa empat topik yang mendominasi pokok bahasan dalam konten yang diunggah di *youtube* pada kategori gaming.

IV. HASIL DAN PEMBAHASAN

A. Studi Literatur

Kami mengambil data mengenai video yang diunggah oleh 10 *youtuber gaming* Indonesia dengan *subscribers* terbanyak. Dengan subscriber yang cukup besar menandakan *youtuber* tersebut memang sangat dikenal dan diminati masyarakat Indonesia khususnya pada kategori *gaming*.

Kami mencatat *channel youtube* milik masing-masing *youtuber* yang mana akan digunakan pada tahap *data crawling*. Pada tabel 1 dan 2 berikut merupakan data *channel*, jumlah subscriber, dan banyaknya video yang diambil pada 28 September 2021.

TABEL I
SEPULUH CHANNEL YOUTUBE GAMING INDONESIA DENGAN JUMLAH
SUBSCRIBERS TERBANYAK

No.	Channel	Subscribers (Juta)
1	Jess No Limit	23,3
2	Frost Diamond	19,6
3	Miawaug	15,7
4	Dylan Pross	15,1
5	Budi Gaming	11,8
6	Frontal	11,5
7	Letda Hyper	9,04
8	Rendy	8,96
9	Kemas	8,27
10	Erpan1140	7,73

B. Data Crawling

Pengambilan data dilakukan terhadap *youtuber gaming indonesia* yang memiliki subscriber terbanyak. Berdasarkan kriteria tersebut, maka diambil sepuluh sampel sebagai bahan penelitian. Daftar *channel* yang diambil ditunjukkan pada Tabel 1 dimana subscriber dari *youtuber gaming indonesia* ada yang telah mencapai hingga 23,3 juta *subscribers* yang dimiliki oleh Jess No Limit. Sedangkan jumlah yang paling sedikit, yakni sejumlah 7,3 ribu *subscribers* dan dimiliki oleh Erpan1140. Hasil yang diperoleh selanjutnya akan dimanfaatkan dalam proses LDA pada tahap selanjutnya.

TABEL III
 JUMLAH VIDEO DARI 10 CHANNEL YOUTUBE GAMING INDONESIA

No.	Channel	Video
1	Jess No Limit	1618
2	Frost Diamond	1731
3	Miawaug	2919
4	Dylan Pross	1482
5	Budi Gaming	229
6	Frontal	74
7	Letda Hyper	699
8	Rendy	1684
9	Kemas	1030
10	Erpan1140	1291
	Total	12757

Setelah *generate API key* dan memperoleh *channel id*, maka dilakukan *crawling* data dan didapatkan judul serta deskripsi yang berisi waktu pengunggahan video dan keterangan dari setiap akun *youtuber* tersebut. Setelah data tersebut didapatkan kemudian disimpan dalam *file* berekstensi *.csv* sebagai data untuk dilakukan proses analisis selanjutnya dan diunggah melalui akun Zenodo [19]. Dari data yang diperoleh maka didapatkan jumlah video dari masing-masing akun *youtuber Gaming* yang dapat dilihat pada Tabel 2. Hasil yang didapatkan menunjukkan setiap *channel* memiliki jumlah video yang beragam, mulai dari 74 hingga mencapai 2919 video yang telah diunggah. Jumlah keseluruhan video dari sepuluh saluran tersebut sebanyak 12757 video.

C. Data Preparation

Setelah dilakukan *data crawling* untuk 10 *channel youtube* yang mana menghasilkan output berupa 10 *file .csv* yang berisi waktu *upload*, judul, durasi, dan deskripsi setiap video pada *channel youtube* tersebut. Kami melakukan *data preparation* untuk menyiapkan data untuk proses LDA.

Pada proses LDA nantinya akan membutuhkan data judul dari setiap video yang diupload oleh setiap *youtuber*. Total dari seluruh video adalah sebanyak 12.757 video. Karena cukup banyak, kami memutuskan untuk menyatukan terlebih dahulu 10 *file .csv* dalam satu *file excel* kemudian menyatukan judul

video secara keseluruhan ke dalam satu dokumen memastikan setiap judul terpisah satu baris antara satu dengan lainnya.

Setelah seluruh judul video telah disatukan, kami pindahkan seluruh judul tersebut ke dalam program *python* yang ada di *Zenodo* [19]. Kami juga menuliskan cukup banyak *stopword* untuk dihilangkan. *Stopword* merupakan kata-kata yang tidak informatif sebagai *keyword* dari topik dalam penelitian ini. Dalam proses LDA nantinya kami juga melakukan *trial and error* untuk menghilangkan *stopword* untuk memastikan *keyword* dalam suatu topik dominan adalah bukan merupakan kata yang tidak relevan.

D. Penerapan LDA

Setelah data disiapkan, maka berikutnya adalah memasukkannya ke proses LDA. Fungsi *Corpora* digunakan untuk memetakan setiap kata, dan mengubahnya menjadi *term dictionary*. Langkah selanjutnya adalah pembuatan *document-term matrix* dengan menggunakan modul *gensim*. Untuk mempermudah proses pemrograman, proses LDA dilakukan dengan memanfaatkan paket *LdaModel*. Dengan demikian terbentuklah model LDA yang kemudian divisualisasikan dalam *bar chart*.

Untuk mengevaluasi model, digunakan dua parameter yakni *perplexity* dan *coherence*. Nilai *coherence* didapatkan dari perhitungan kemiripan semantik antar kata dalam satu topik. Maka semakin tinggi nilai koherensi, semakin baik model LDA yang dibuat. Sedangkan *perplexity* menilai ‘tingkat keterkejutan’ suatu model atas data yang diprosesnya. *Perplexity* dihitung dengan *normalized log-likelihood*, sehingga semakin rendah nilai tersebut, maka semakin baik model yang dibuat.

E. Analisis

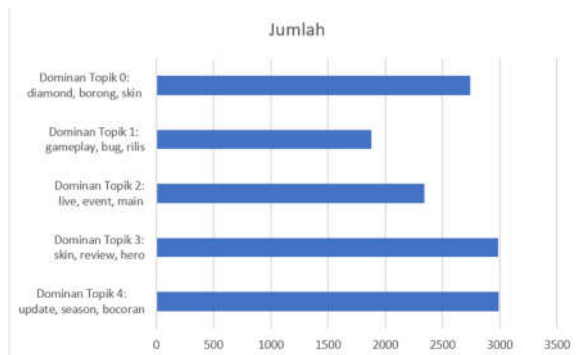
Setelah melakukan proses LDA pada data, selanjutnya dilakukan analisis terhadap data tersebut. Untuk menentukan jumlah topik yang optimal yakni dengan melihat parameter nilai *perplexity* dan nilai *topic coherence*. Semakin rendah nilai *perplexity* maka semakin baik modelnya. Sebaliknya semakin tinggi nilai *topic coherence* semakin bagus model LDA yang digunakan. Dengan melakukan percobaan perhitungan dengan jumlah topik sebanyak 3, 4 dan 5 didapatkan nilai *perplexity* dan nilai *topic coherence* yang ditampilkan dalam Tabel 3.

TABEL IIIII
 NILAI PERPLEXITY DAN TOPIC COHERENCE

Jumlah Topik	Nilai Perplexity	Topic Coherence
3	-40.06654639351327	0.44490896742473857
4	-40.14703894612008	0.47179459270522410
5	-40.17842609602861	0.47420506316078220

Dari hasil yang didapatkan menunjukkan bahwa nilai yang ada hanya memiliki selisih yang kecil. Dengan melihat dari nilai *perplexity* sebagai acuan dalam menentukan jumlah topik yang baik maka dipilihlah jumlah topik sebanyak 5 yang menunjukkan nilai *perplexity* paling rendah. Selain itu pada 5

topik juga memiliki nilai *topic coherence* yang tertinggi. Selanjutnya dihasilkan diagram yang berisi topik yang paling koheren yang ditunjukkan pada Gambar 2.



Gbr. 2 Grafik jumlah video pada 5 topik yang dominan beserta keyword yang berkorelasi pada setiap topik

Untuk melihat data secara lebih jelas dan detail dapat dilihat melalui Tabel 4 yang menunjukkan jenis topik beserta jumlahnya. Setelah melalui proses LDA dan analisis data maka didapatkan daftar topik yang dominan dari para *youtuber Gaming Indonesia* dengan *subscribers* terbanyak.

TABEL IVV
JUMLAH VIDEO PADA 5 TOPIK YANG DOMINAN BESERTA KEYWORD YANG BERKORELASI PADA SETIAP TOPIK

No	Dominan_Topic	Keywords	Jumlah
1	0	update, season, bocoran	2990
2	1	skin, review, hero	2987
3	2	live, event, main	2340
4	3	gameplay, bug, rilis	1879
5	4	diamond, borong, skin	2741

Dari Tabel 4 dan Gambar 2 menunjukkan hasil persebaran topik ditampilkan dalam 3 kata di setiap topiknya. Pada topik dominan pertama membahas bocoran *update* dan *season* baru pada *game* online yang dimainkan. Topik dominan pertama merupakan topik yang paling banyak muncul dengan jumlah 2990 video. Topik dominan kedua melakukan *review hero* dan *skin* baru yang ada dalam *game*. Topik dominan yang kedua juga cukup banyak muncul, yaitu sebanyak 2987 video.

Topik dominan ketiga yaitu konten *live* bermain *game* disaat ada *event game* yang menarik. Topik dominan ketiga muncul sebanyak 2340 video. Topik dominan keempat yaitu membahas *bug* yang terjadi pada *patch* baru yang rilis, yang mana dengan adanya *patch* dan *bug* tersebut akan mengubah *gameplay* seseorang. Topik dominan keempat adalah yang paling sedikit muncul, yaitu sebanyak 1879 video. Topik dominan kelima yaitu membongkar *skin* dalam *game* menggunakan *diamond* yang cukup banyak. Topik dominan kelima muncul pada 2741 video.

V. KESIMPULAN

Berdasarkan analisis output yang dilakukan pada metode LDA, kami memutuskan menggunakan 5 topik dominan dengan pertimbangan nilai *persplexity* dan *topic coherence*. Dalam 5 topik dominan tersebut mengandung tiga *keyword* yang berkorelasi pada setiap topiknya.

Pada topik dominan pertama membahas bocoran *update* dan *season* baru pada *game* online yang dimainkan. Topik dominan kedua melakukan *review hero* dan *skin* baru yang ada dalam *game*. Topik dominan ketiga yaitu konten *live* bermain *game* disaat ada *event game* yang menarik. Topik dominan keempat yaitu membahas *bug* yang terjadi pada *patch* baru yang rilis, yang mana dengan adanya *patch* dan *bug* tersebut akan mengubah *gameplay* seseorang. Topik dominan kelima yaitu membongkar *skin* dalam *game* menggunakan *diamond* yang cukup banyak.

Dari hasil tersebut, ditunjukkan bahwa konten dominan yang dibuat 10 *youtuber* terbesar di Indonesia berkaitan dengan 5 topik dominan tersebut, maka dari itu dapat menjadi referensi bagi seseorang yang ingin menjadi *youtuber gaming* atau *youtuber gaming* yang baru memulai untuk membuat konten mengenai topik-topik yang dominan tersebut. Diharapkan dengan mengikuti konten tersebut, seorang *youtuber gaming* dapat memperoleh lebih banyak penonton dan *subscriber*.

Ada beberapa hal yang perlu dilakukan perbaikan sehingga dapat bermanfaat untuk penelitian masa depan. Rinciannya dijelaskan sebagai berikut: Disarankan bagi peneliti selanjutnya untuk menemukan metode LDA yang lebih efektif, yaitu dengan langsung mengimpor *file* tanpa harus memasukkan keseluruhan isi data pada saat melakukan proses LDA. Aktivitas perayapan data memerlukan tingkat keberhasilan pengiriman paket yang konsisten sehingga diperlukan fasilitas internet yang memadai. Kegagalan ini dapat menyebabkan batas waktu koneksi sehingga proses *crawling* data perlu diulang.

REFERENSI

- [1] Billah, M. F. (2021). Pengaruh *Game Mobile Legend* Terhadap Perkembangan Emosi Siswa Mi Al Hidayah Tarik Sidoarjo.
- [2] Campbell, J. C., Hindle, A., & Stroulia, E. (2014). Latent Dirichlet Allocation: Extracting Topics from Software Engineering Data.
- [3] Cayari, C. "The *Youtube* effect: How *Youtube* has provided new ways to consume, create, and share music," *International Journal of Education & the Arts*, vol. 12, hal. 6, 2011.
- [4] Ginting, B. S., & Ramadhan, F. "Perancangan *Game* Become a King Berbasis Artificial Intelligence," *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, vol. 2, hal. 12-21, 2018.
- [5] Golnari, G., Li, Y., & Zhang, Z. (2014). What Drives the Growth of *Youtube*? Measuring and Analyzing the Evolution Dynamics of *Youtube* Video Uploads.
- [6] Haryanto, H. "Reward Dinamis dalam Skenario Adaptif Menggunakan Metode Finite State Machine pada *Game* Edukasi," *Journal of Applied Intelligent System*, vol. 1, hal. 144-153, 2016.
- [7] Hosch, W. L. (2021, September 15). *Youtube*. (2021) Encyclopedia Britannica. [Online], <https://www.britannica.com/topic/Youtube>, tanggal akses: 22 September 2021

- [8] Indrayuni, E. (2019). Klasifikasi Text Mining Review produk kosmetik Untuk Teks Bahasa Indonesia Menggunakan ALGORITMA naive Bayes. *Jurnal Khatulistiwa Informatika*, vol. 7, hal. 1, 2019.
- [9] Iqbal, M. (2020). *Youtube Revenue and Usage Statistics* (2020). Retrieved from *Business of Apps*. [Online], <https://www.businessofapps.com/data/Youtube-statistics/>, tanggal akses: 22 September 2021.
- [10] Jayani, D. H. (2020). Orang Indonesia Habiskan Hampir 8 Jam untuk Berinternet. <https://databoks.katadata.co.id/datapublish/2020/02/26/indonesia-habiskan-hampir-8-jam-untuk-berinternet>, tanggal akses: 22 September 2021.
- [11] Kresnawan, H., Felle, S. G., Mokay, H. G., & Rakhmawati, N. A., "Analyzing Main Topics Regarding The Electronic Information and Transaction Act in Instagram Using Latent Dirichlet Allocation," *Data Science: Journal of Computing and Applied Informatics*, vol. 5, no. 2, hal. 71-84, 2021.
- [12] Mellyaningsih, A., "Motif Subscriber Menonton *Channel Youtube* Raditya Dika," *Jurnal E-Komunikasi Program Studi Ilmu Komunikasi Universitas Kristen Petra, Surabaya*. vol. 4, no. 1, 2017.
- [13] Nurazizah, R. S. (2017). Analisis Resepsi pada *Channel Game* Reza Oktovian. Retrieved 2021.
- [14] Prakerti, A. I, et al. "Model Latent Dirichlet Allocation Pada Perilaku Siswa Menggunakan Media Pembelajaran Daring," *INFORMATION MANAGEMENT FOR EDUCATORS AND PROFESSIONALS : Journal of Information Management*, vol. 5, hal. 35-44, Dec. 2020.
- [15] Rahadian, M. F., Suyatno, A., & Maharani, S. (2017). Penerapan Metode Finite State Machine Pada *Game* "The Relationship".
- [16] Rakhmawati, A., Nikmah, N. L., Perwira, R. D. A., & Rakhmawati, N. A., "Analisis topik konten *channel Youtube* K-pop Indonesia menggunakan Latent Dirichlet Allocation," *Jurnal Ilmiah Sistem Informasi*, vol. 11, hal. 16-25, Jan 2021.
- [17] Social Blade LLC. (2021). Jess No Limit's real-time subscriber count - socialblade. *Social Blade* [online], <https://socialblade.com/YouTube/c/jessnolimit/realtime>, tanggal akses: 21 September 2021.
- [18] Suhartono, D. (2018). Latent Dirichlet Allocation (LDA). <https://socs.binus.ac.id/2018/11/29/latent-dirichlet-allocation-lda/>, tanggal akses: 22 September 2021.
- [19] Waskitho, R., Nuha, M., Rahman, D. and Rakhmawati, N., 2021. Dataset 10 Youtuber Gaming Indonesia [online], DOI: 10.5281/zenodo.5536343, tanggal akses : 21 September 2021.
- [20] Y. Du, Y. Yi, X. Li et al. (2020) Extracting and tracking hot topics of micro-blogs based on improved Latent Dirichlet Allocation, 87.