

A Survey on Big Data in the Media and Entertainment Industry

Annisaa Nurhayati

Sekolah Tinggi Elektro dan Informatika
Institut Teknologi Bandung
Bandung, Indonesia
annisaa.nurhayati@student.itb.ac.id

Abstract—Big Data has affected all industries, including the media dan entertainment industries. The popularity of using mobile devices and the internet has changed the way people enjoy entertainment. This popularity also generates data streams from many sources with various data formats and large volumes, known as big data. Carrying out big data analysis can help the media industry and entertainment achieve its goals, like providing content that makes users happy, provides user experience, and increases profits. Many researchers have conducted research on the use of big data in the media and entertainment industries. The purpose of this paper is to provide an overview of the problems, challenges and various technologies related to Big Data in the media and entertainment industries.

Keywords—big data, media, entertainment

I. PENDAHULUAN

Penggunaan internet dan perangkat *mobile* yang semakin populer telah mengubah cara masyarakat dalam menikmati hiburan. Kepopuleran ini berdampak pada meningkatnya data yang berjumlah besar dan memiliki berbagai format, dikenal dengan istilah Big Data.

Big Data merupakan aset informasi yang memiliki volume besar, kecepatan tinggi, dan/atau banyak format, yang membutuhkan pemrosesan baru untuk meningkatkan pembuatan keputusan, penemuan wawasan, dan optimisasi proses[1]. Tren Big Data ini telah mempengaruhi semua industri, tidak terkecuali industri media dan hiburan.

Pada industri media, Big Data didefinisikan sebagai 4V, yaitu *volume* (ukuran) data, *velocity* (kecepatan) data, maksudnya data yang butuh dianalisis secepatnya (terutama berita) dalam *variety* (beragam) format data, baik yang terstruktur maupun data tidak terstruktur yang semakin meningkat jumlahnya, yang seluruhnya memiliki *value* (nilai) yang berpotensi dalam memberikan jurnalistik kualitas tinggi serta wawasan dan pendapatan bisnis [1].

Big Data memungkinkan industri media mencapai tujuannya, seperti menyesuaikan konten untuk pelanggan dan membantu perusahaan media mendapatkan wawasan tentang perilaku pelanggan. Data yang besar ini dapat menjadi kekuatan besar dalam penguatan komersial. Untuk industri media, big data berperan penting dalam rangka mendapatkan keuntungan bisnis yang tinggi. Konten media tradisional disediakan menggunakan format terbatas, namun kini telah digantikan oleh layanan multimedia yang dibayar per tampilan, siaran sesuai permintaan, siaran langsung, dan lain-lain [2].

Dengan bantuan Big Data, perusahaan media dan hiburan dapat mengetahui mengapa pelanggan tetap berlangganan atau membatalkan langganan, mengetahui program apa yang pelanggan sukai, juga mendapatkan wawasan secara rinci tentang harga dan respon langganan. Data yang besar juga dapat memberikan informasi tentang genre hiburan apa yang sedang dinikmati oleh penonton secara *real-time*. Dengan mengetahui preferensi pelanggan, dapat membantu industri media dalam menyusun strategi iklan. Hal ini dapat membantu pencapaian tujuan, misalnya untuk memperoleh rating televisi yang tinggi [2].

Pada industri media seperti surat kabar, televisi, majalah, dan penerbitan *online*, Big Data dapat memberikan analisis pangsa pasar untuk mendapatkan pemahaman dan penargetan pelanggan yang lebih baik, sebagai alat untuk memahami database publik dan pribadi untuk penceritaan jurnalistik, sebagai alat untuk mengelola dan mencari jumlah video, media sosial, dan konten lainnya yang sedang digandrungi penonton, sebagai alat untuk menargetkan iklan dan kampanye iklan, sebagai alat untuk mengotomatisasi produksi cerita teks dan video, sebagai alat untuk mengidentifikasi pemborosan dan memungkinkan efisiensi, dan banyak lagi [3].

Industri media dan hiburan seringkali berada di garis terdepan dalam mengadopsi teknologi terbaru. Permasalahan utama yang mendorong perusahaan media memanfaatkan kemampuan big data adalah kebutuhan untuk mengurangi biaya operasional dalam dunia bisnis yang semakin kompetitif, dan pada saat yang sama, menghasilkan keuntungan dari penyampaian konten dan data melalui beragam platform dan produk [3].

Ada tiga area utama dimana Big Data berpotensi mempengaruhi keadaan yang ada dan memstimulasi pertumbuhan ekonomi pada sektor media dan hiburan [3]:

- Produk dan Layanan: Big Data mendorong bisnis media untuk memiliki kemampuan dalam menerbitkan konten dalam bentuk yang lebih mutakhir.
- Pelanggan dan Pemasok: Perusahaan media yang ambisius akan menggunakan Big Data untuk mencari tahu lebih banyak informasinya, seperti pilihan yang disukai, profil, dan perilaku. Informasi tersebut akan digunakan untuk membangun hubungan yang lebih mengikat dengan pelanggannya.
- Infrastruktur dan Proses: Proses dan budaya organisasi juga butuh menyesuaikan dengan ekspektasi yang ditawarkan Big Data. Kegagalan dalam mengubah budaya dan keterampilan pegawai dapat berpengaruh pada perusahaan yang berpeluang memberikan keuntungan namun tidak dapat beradaptasi pada model bisnis berdasarkan data.

Tujuan utama dari survei ini untuk memperlihatkan berbagai penelitian terkini mengenai pemanfaatan Big Data pada industri media dan hiburan. Survei ini menyajikan tinjauan mengenai teknik, tantangan, dan solusi Big Data pada industri media dan hiburan, serta arahan yang sesuai untuk penelitian di masa mendatang. Target dari survei ini adalah penelitian-penelitian mengenai teori dan teknologi yang memberikan kemajuan dan kebaruan di industri media dan hiburan dengan menggunakan Big Data. Dengan kata lain, survei ini bertujuan untuk menjembatani kesenjangan antara tantangan pada industri media dan hiburan dengan solusi Big Data yang ada saat ini.

Pada bagian selanjutnya, yaitu bagian II akan membahas penelitian-penelitian terkini dalam penggunaan Big Data di industri media dan hiburan. Dalam bagian III, beberapa tantangan dan peluang berdasarkan hasil survei dibahas. Terakhir, pada bagian IV menyimpulkan keseluruhan dari hasil penelitian ini dan mengeksplorasi beberapa arahan dan pandangan untuk penelitian di masa mendatang.

II. STUDI LITERATUR SURVEI

Ada banyak tantangan dan peluang analisis Big Data pada industri media dan hiburan, seperti yang dapat dilihat pada Tabel I. Oleh karena itu, banyak penelitian-penelitian yang berfokus untuk menjawab tantangan ini. Pada makalah survei ini, penelitian-penelitian tersebut dibahas dan dikategorikan berdasarkan hasil akhir penelitiannya. Perbandingan antara penelitian-penelitian ini dapat dilihat juga pada Tabel II.

TABEL I. TANTANGAN BISNIS DAN PELUANG PADA PERUSAHAAN MEDIA DAN HIBURAN [4]

Area Fungsional	Tantangan Bisnis	Peluang
Pemasaran Media dan Hiburan	Mendapatkan keuntungan besar per dolar dari pengeluaran pemasaran	<ul style="list-style-type: none"> • Meningkatkan pangsa pasar • Meningkatkan citra dan loyalitas merek • Meningkatkan transaksi lintas saluran • Meningkatkan respon iklan/promosi
Pendapatan periklanan	Menunjukkan nilai media dan hiburan kepada pengiklan yang ada dan potensial	<ul style="list-style-type: none"> • Meningkatkan volume dan pendapatan iklan • Meningkatkan nilai iklan • Mempertahankan pengiklan dengan pindah ke saluran yang tepat
Pembuatan konten, akuisisi, dan pemrograman	Membuat atau membeli dan menjadwalkan konten yang paling menarik yang diinginkan oleh pelanggan untuk pendistribusian yang efisien	<ul style="list-style-type: none"> • Memberikan pengalaman menonton yang berbeda dan dipersonalisasi • Meningkatkan jumlah penonton • Menawarkan konten yang menguntungkan • Meningkatkan minat pengiklan • Menjadwalkan/membuat konten tersedia melalui saluran yang paling relevan
Pengarsipan dan penarikan konten	Mengarsipkan, menarik, atau menjual konten dan penawaran ketika tidak lagi menguntungkan	<ul style="list-style-type: none"> • Mempertahankan penonton/pelanggan yang telah didapat • Mengganti konten lama dengan konten yang lebih menarik • Mempertahankan/menumbuhkan pendapatan iklan • Meningkatkan minat pengiklan • Mengelola biaya infrastruktur (misal, penyimpanan)
Pengelolaan tempat hiburan (misal taman hiburan)	Mengoptimalkan pengalaman pelanggan sambil meningkatkan margin	<ul style="list-style-type: none"> • Mengarahkan pelanggan secara efektif • Mengoptimalkan pengalaman tempat duduk/video/audio • Meningkatkan/mengganti penawaran yang memiliki kinerja buruk • Penawaran cross-sell/up-sell • Meningkatkan kehadiran • Meraih respon positif

		<ul style="list-style-type: none"> • Mengoptimalkan staf, fasilitas, kesediaan makanan/minuman, dll
--	--	--------------------------------------------------------------------------------------------------------------------

A. Analisis Tren

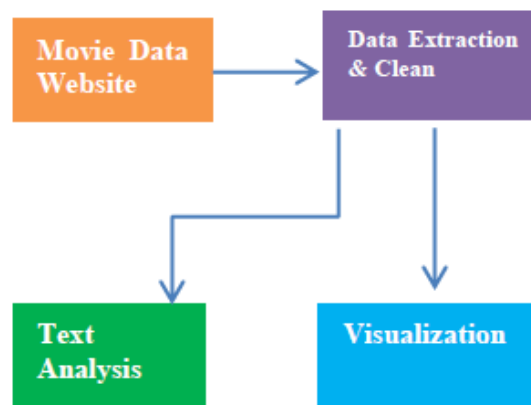
Penelitian [2] menyusun analisis data dari industri media dan hiburan serta peran penting yang Big Data mainkan pada industri ini. Tahapan proses yang dilakukan dalam penelitian ini meliputi:

- Memprediksi kebutuhan pelanggan
- Mengoptimalkan penjadwalan
- Meningkatkan perolehan dan penyimpanan
- Penargetan iklan
- Monetisasi konten dan pengembangan konten baru

Untuk menghasilkan generasi terbaik di industri media dan hiburan, sangat penting bagi industri dengan bantuan analisis big data untuk memprediksi pilihan pelanggan dan menggunakan data tersebut untuk merekomendasikan konten yang disukai agar pelanggan tetap puas dan bahagia [2].

Untuk penelitian lebih lanjut, Big Data dapat digunakan untuk menganalisis konten pengguna berdasarkan episode sebuah pertunjukan/program yang selalu ditonton berulang kali. Hal ini dapat digunakan untuk mengetahui program tertentu yang sangat disukai pelanggan sehingga dapat menyediakan layanan yang dapat membuat pelanggan terikat pada program. Dengan bantuan Big Data, analisis data pengguna pada platform musik dapat digunakan untuk memberikan rekomendasi pada pengguna tentang artis dan genre musik, serta untuk improvisasi proses bisnis dengan memantau dan menganalisis aktivitas sehari-hari pelanggan [2].

Penelitian [5] dan [6] berfokus pada industri film di China. Penelitian [5] mencari aturan dari pasar film *box office* berdasarkan data historis film yang dirilis di China. Mendapatkan aturan pasar film akan membantu pembuat film mengerti apa yang disukai penonton dan membuat keputusan untuk proyek film selanjutnya. Pada penelitian ini dianalisis secara visual tren sepanjang tahun dan bulan pada film *box office* dan proporsi film *box office* berdasarkan genre dan bahasa. Selain itu, dianalisis juga nilai *rating* film, untuk mengetahui kualitas film. Hal pertama yang dilakukan pada penelitian ini adalah menggabungkan data film *box office* yang telah dirilis di China pada 15 tahun terakhir dan informasi dasar (seperti data rilis, bahasa, genre, negara, dsb.), selain itu penelitian ini juga mengumpulkan komentar dan respon penonton film yang jumlahnya mencapai 100.000 komentar. Data yang telah diperoleh kemudian dianalisis pendistribusian data film dalam bulan, tahun, genre, dan bahasa yang digunakan.

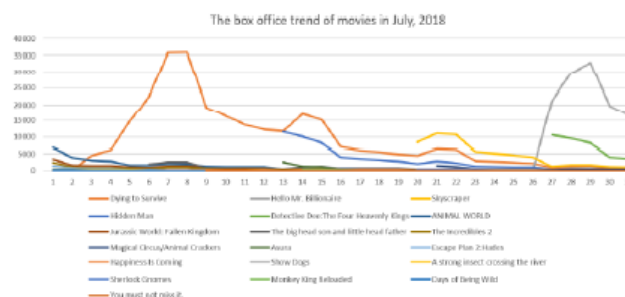


Gambar 1. Proses *data mining* [5]

Sementara, penelitian [6] mengusulkan sistem analisis visual untuk big data film *box-office* di China, yang meliputi tiga bagian, yaitu gabungan data *real-time*, data visualisasi, dan data analisis.



Gambar 2. Contoh sistem display *real-time box office* [6]



Gambar 3. Tren film *box office* rilis Juli 2018 [6]

Dari sudut pandang ISP (*Internet Service Provider*), peningkatan popularitas dari perangkat *mobile* dan internet memberikan tantangan baru seperti persaingan yang beragam dan perilaku pelanggan yang berubah-ubah. Oleh karena itu untuk merespon permintaan pelanggan yang berubah, penggunaan Big Data untuk menganalisis pelanggan yang sudah ada dan pelanggan yang potensial menjadi tren dalam bisnis perancangan rencana marketing dan produk. Penelitian [7] bertujuan untuk mencari pelanggan potensial dari layanan video *online* Chunghwa Telecom dengan mempelajari perilaku pengguna pada beberapa aplikasi video. Hasil dari penelitian ini akan digunakan oleh divisi marketing untuk merancang rencana promosi. Perusahaan butuh untuk mengumpulkan dan menganalisis data pelanggan untuk mengidentifikasi karakteristik target pelanggan.

Penelitian [8] menganalisis perilaku pengguna berdasarkan perubahan daftar putar musik yang didengarkan, yang diekstraksi dengan mengaplikasikan Hidden Markov Model (HMM). HMM digunakan untuk membuat model perilaku pengguna. Hasilnya, ditemukan sedikit perubahan yang baik dalam daftar putar musik yang didengarkan bagi pengguna. Penelitian ini juga menganalisis perilaku pengguna saat perubahan baik itu terjadi. Hasilnya, pengguna cenderung merubah daftar putar musik mereka saat sedang menempuh perjalanan dan memilih tidak mengubah daftar putar musik mereka saat tengah malam. Sebagai tambahan, hubungan antara tren musik yang didengarkan pengguna dan gaya hidup pengguna pun dianalisis.

B. Rekomendasi Konten

Penelitian [9] mengusulkan sistem rekomendasi video pribadi dengan bantuan *cloud* berdasarkan pembelajaran terdistribusi online. Pertumbuhan pesat layanan multimedia dan banyaknya tawaran konten video di jejaring sosial online membuat pengguna kesulitan dalam mendapatkan konten sesuai minat mereka, oleh karena itu beragam sistem rekomendasi yang dipersonalisasi telah diusulkan. Namun, percepatan penyebaran data di media sosial telah menuntun ke era Big Data, yang dapat menghambat proses rekomendasi video. Selain itu, tidak ada dari sistem rekomendasi yang mempertimbangkan konteks privasi pengguna (misal, status sosial, usia, dan hobi) dan repositori video milik penyedia layanan, yang sangat sensitif dan memiliki nilai komersial yang signifikan. Pada kerangka kerja yang diusulkan, video direkomendasikan disesuaikan dengan konteks pengguna, selagi

secara simultan mengadaptasi strategi pemilihan video berdasarkan respon klik pengguna untuk memaksimalkan jumlah klik pengguna.

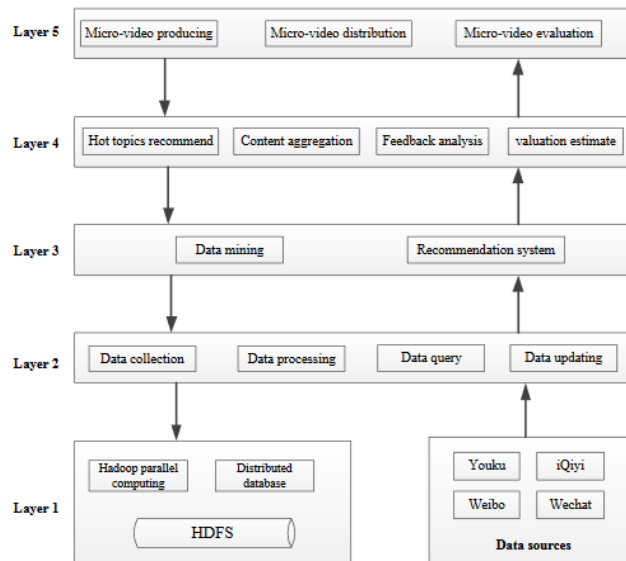
Penelitian [10] mengusulkan dan memeriksa pendekatan baru untuk menghasilkan *latent embeddings* untuk item podcast dengan memanfaatkan informasi yang telah dikumpulkan dari seluruh fitur berbasis teks yang terkait dengan item audio. *Embeddings* yang dihasilkan menggunakan teknik *Natural Language Processing* (NLP) pada item podcast ini dapat digunakan untuk mengukur atau mengindikasikan kemiripan konten antara berbagai item podcast.

Podcast merupakan kombinasi antara aktivitas saluran radio FM dengan kiriman blog ekonomis. Podcast sangat nyaman dan mudah jika memiliki kemampuan dan konektivitas internet terbatas. Sebagai contoh, sejumlah besar podcast dinikmati saat sedang berada di dalam mobil, gym, dll. Ketika volume dan keberagaman konten sangat besar dengan berbagai sumber dari musik, buku audio, pertunjukan komedi, berita, dan sebagainya, sehingga menjadi sulit untuk mengkategorikan secara manual atau menandai item audio, agar lebih mudah ditemukan pengguna. Selain itu, ketidaklengkapan metadata pada audio seperti tanda ID3 (genre, artis, dll) menyebabkan kekurangan fitur untuk sistem rekomendasi untuk mempelajari kemiripan item (berbasis konten) atau *latent embeddings* dari pengguna yang memiliki pemikiran sama (penyaringan kolaboratif) untuk membantu pengguna mengeksplorasi item yang berpotensi menarik minat [10].

Sementara itu, penelitian [11] mengusulkan pendekatan terbaru dalam perekomendasi buku, dengan memanfaatkan big data yang dihasilkan dari ribuan pengguna situs katalog buku sosial dan menjadikannya sebagai meta-anotasi buku. Dengan digitalisasi buku dalam skala besar, banyak strategi lama dalam penemuan buku, yang dirancang untuk ruang fisik dengan jumlah item terbatas, tidak berfungsi dengan baik dalam ruang digital. Ada pertumbuhan penelitian yang berfokus dalam pengembangan pendekatan inovatif untuk pencarian dan penelusuran biasanya menggunakan sistem rekomendasi.

Setelah mempelajari kemiripan semantik diantara banyaknya koleksi buku dengan menerapkan algoritma *natural language processing* (model topik probabilistik dan jaringan saraf semantik) ke meta-anotasi teks simbolis, penelitian ini menyajikan bagaimana mengkonstruksi mesin rekomendasi semantik yang lebih tepat dan deskriptif dan memperkaya mesin rekomendasi berbasis konten [11].

Penelitian [12] mengusulkan sebuah sistem rekomendasi video mikro. Video mikro adalah format terbaru dari media informasi. Video mikro merupakan video berdurasi pendek, yang berlangsung selama 30 detik hingga 300 detik. Video mikro berdurasi singkat ini sangat populer dikalangan anak muda, karena para remaja lebih suka menonton video mikro pada waktu senggang mereka melalui perangkat *mobile*. Permasalahan yang dihadapi pembuat video mikro adalah mereka tidak mengetahui berapa banyak orang yang menyukai video mereka, dan tidak tahu sudah berapa kali video mereka ditonton. Oleh karena itu, penelitian ini mengusulkan sistem rekomendasi video mikro (MRS). Salah satu tujuan MRS adalah sebagai ikhtisar video mikro bagi produser. Dengan cara ini, produser tahu berapa banyak pengguna menyukai video mereka dan berapa kali video mereka sesuai dengan permintaan. Tujuan lain dari MRS ditujukan untuk pengguna. Sistem dapat menganalisis video favorit pengguna dan riwayat menonton, yang secara otomatis akan memberikan video yang sesuai kepada pengguna. Penelitian ini menggunakan kerangka kerja Hadoop sebagai platform analisis big data secara offline dan menggunakan Mahout untuk memproses data.

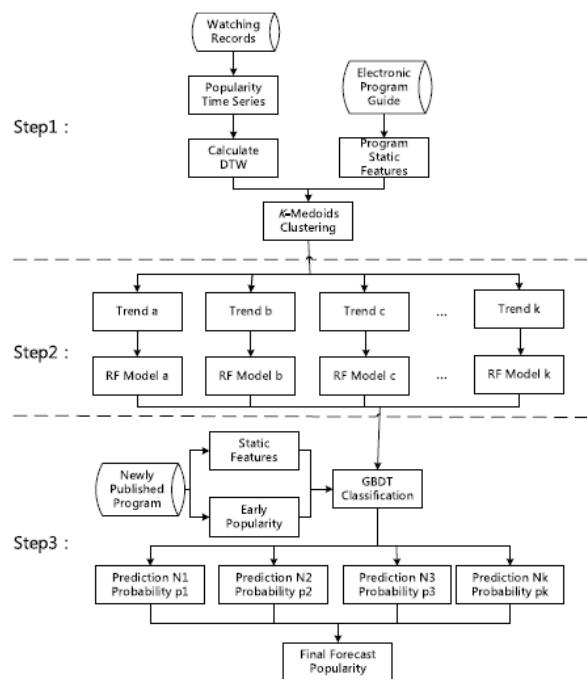


Gambar 4. Sistem rekomendasi video mikro [12]

Rekomendasi jenis hiburan yang dipersonalisasi diperlukan untuk membantu jutaan orang dalam mempersempit jenis hiburan potensial yang sesuai dengan minatnya. Layanan rekomendasi ini biasanya tergantung pada algoritma *machine-learning*, yang memecah jenis hiburan menjadi daftar panjang, lalu menghubungkan dan mencocokkan elemen-elemen ini dengan preferensi pengguna. Penelitian [13] mengusulkan algoritma rekomendasi hiburan yang merekomendasikan jenis hiburan pada sekelompok pengguna. Algoritma ini tidak hanya bergantung pada peringkat tapi juga waktu dari peringkat. Algoritma ini juga merekomendasikan jenis hiburan populer pada produser.

C. Prediksi Konten dan Minat Pengguna

Penelitian [14] menyajikan pendekatan untuk meningkatkan prediksi berbasis deteksi tren. Prediksi popularitas program yang tepat guna dan tepat waktu merupakan nilai yang bagus bagi penyedia konten, pengiklan, dan penyelenggara siaran televisi. Informasi ini memberikan keuntungan untuk penyelenggara program televisi dan dapat membantu pengiklan untuk memformulasikan rencana investasi iklan yang sesuai. Dalam hal teknis, prediksi popularitas program yang tepat dapat mengoptimalkan sistem penyiaran, seperti strategi jaringan penyampaian konten dan strategi *cache*. Metode prediksi yang sudah ada biasanya meminta jumlah sampel yang besar dan waktu pengujian yang lama, namun memberikan keakuratan prediksi yang kurang tepat untuk program yang memiliki kenaikan atau penurunan popularitas drastis.



Gambar 5. Metode prediksi popularitas penyiaran program TV [14]

Langkah pertama yang dilakukan adalah mendeteksi tren perubahan popularitas. Menghitung *Dynamic Time Warping-distance* (DTW) pada *historical-record time series*. Langkah kedua, mengidentifikasi tren penyiaran dan memisahkannya ke dalam beberapa tipe. Untuk tren penyiaran TV, pengklusteran berdasarkan *time-series* dilakukan menggunakan DTW berbasis K-medoids. Langkah ketiga, membangun model prediksi trend-spesifik menggunakan regresi *random forest*. *View record* dibagi menjadi 4 grup berdasarkan tren dan memasukan ke dalam RF model bersama dengan fitur statis. Terakhir, menggunakan gradient boosting decision tree (GBDT) untuk mengklasifikasikan popularitas *time-series* dari program yang baru dipublikasikan ke dalam tren dan mendapatkan hasil prediksi akhir berdasarkan nilai prediksi dari 4 model dan kemungkinan klasifikasi.

D. Platform Media dan Hiburan

Penelitian [15] mengembangkan sebuah platform big data terpadu untuk analisis sosial TV, mengambil wawasan yang memiliki nilai dari respon tv sosial secara *real-time*. Platform yang dibangun merupakan platform cloud-centric dengan dukungan SDN (*Software Defined Networking*), menyediakan mesin virtual sesuai permintaan dan jaringan yang dapat dikonfigurasi ulang. Arsitektur sistem terdiri dari tiga komponen kunci, yaitu sistem *crawler* data yang kuat, sistem pemrosesan big data menggunakan SDN, dan sistem analisis media sosial. Sistem *crawler* data mengadopsi arsitektur terdistribusi untuk menghindari kendala akses sistem OSN dalam melakukan *crawler* data yang cukup mengenai program TV yang diminati. Sistem pemrosesan big data menggunakan SDN mengintegrasikan SDN dan Hadoop, dan memanfaatkan keuntungan SDN untuk memindahkan data antara unit proses yang berbeda untuk mempercepat laju pemrosesa data. Sistem analisis media sosial mengekstraksi persepsi publik dan pengetahuan yang berhubungan dengan program TV berbasis data mikroblog.

E. Churn Pelanggan

Persaingan pasar media dan hiburan yang semakin ketat menyebabkan pengelolaan *churn* pelanggan menjadi hal yang penting bagi perusahaan. *Churn* pelanggan mengacu pada kehilangan pelanggan yang beralih dari satu perusahaan ke pesaing lain pada periode waktu tertentu. *Churn* pelanggan dapat menyebabkan kerugian ekonomi yang sangat besar dan bahkan dapat merusak citra publik sebuah perusahaan [16].

Penelitian [16] menggunakan Big Data untuk pengelolaan *churn* pelanggan dengan menyediakan metode baru untuk membantu perusahaan dalam mengurangi resiko *churn* pelanggan dan

mendapatkan keuntungan yang lebih tinggi. Metode yang diusulkan merupakan algoritma pengklusteran baru bernama semantic-driven SCM (SDSCM), yaitu metode yang mengintegrasikan AFS (*Axiomatic Fuzzy Sets*) dan SCM (*Subtractive Clustering Method*). SDSCM meningkatkan akurasi pengklusteran SCM dan *k-means*. Selain itu, mengurangi resiko ketidaktepatan manajemen operasi menggunakan AFS.

Penelitian ini juga memodifikasi SDSCM versi awal ke SDSCM big data dan mengimplementasikannya pada kerangka kerja *MapReduce*. Selanjutnya, penelitian ini menyelesaikan masalah *churn* pelanggan di perusahaan telekomunikasi China dengan SDSCM big data dan algoritma *k-means* big data.

Masalah *churn* pelanggan juga dirasakan oleh industri hiburan permainan online. *Free-to-play* dan bisnis berbasis layanan kini mendominasi pasar permainan online, di mana alih-alih membebaskan biaya awal tunggal untuk lisensi permainan, perusahaan menghasilkan uang dengan berlangganan, iklan, atau transaksi mikro dari pemain yang membayar seumur hidup mereka sebagai pelanggan. Dengan melakukan hal ini, pembuat permainan mulai menghadapi masalah yang dihadapi oleh perusahaan dari industri media dan hiburan lainnya, yaitu *churn* pelanggan [17].

Dalam permainan *free-to-play*, dimana tidak ada layanan berlangganan yang menghubungkan antara perusahaan permainan dengan pemain, pemain hanya perlu berhenti bermain game tanpa perlu memberitahu perusahaan tentang keputusannya meninggalkan permainan. Ini cukup sulit bagi perusahaan karena mereka kehilangan kesempatan untuk mengembalikan keputusan para pemain [17].

Pada pasar permainan online, perusahaan hanya memiliki informasi terkait aktivitas pemain dalam permainan mereka. Informasi ini dikumpulkan dan dianalisis untuk mendapatkan data perilaku, hal inilah yang dijadikan perusahaan sebagai dasar dalam memprediksi *churn* pemain. Catatan pemain seperti tanggal dan waktu memulai permainan menjadi data andalan bagi perusahaan. Pemain mengubah perilaku mereka ketika akan meninggalkan permainan, terutama yang berkaitan dengan waktu di antara sesi permainan. Terdapat cara sederhana untuk mengubah data catatan waktu memulai permainan menjadi sebuah fitur, terdapat konsep populer pada literatur marketing untuk menganalisis nilai pelanggan yaitu RFM (*Recency, Frequency, and Monetary*). Pada penelitian [17] dikembangkan sebuah model prediksi *churn* dari data nyata, klasifikasi probabilitas berdasarkan algoritma *k-nearest neighbors* (kNN) menggunakan 4 fitur berbeda untuk catatan login, yaitu : 1) berbasis RFM; 2) domain waktu; 3) domain frekuensi; dan 4) domain bidang frekuensi-waktu.

TABEL II. PERBANDINGAN PENELITIAN BIG DATA PADA INDUSTRI MEDIA DAN HIBURAN

Kategori	Penulis	Tahun	Bidang Minat	Deskripsi
Analisis Tren	Suri dan Singh [2]	2018	<ul style="list-style-type: none"> • Perkiraan data • <i>R Programming</i> • <i>Big data</i> • Industri media dan hiburan • Statistika 	Menyajikan peran signifikan big data pada industri media dan hiburan.
	Cui dkk.[5]	2018	<ul style="list-style-type: none"> • Film yang dirilis di China • Analisis data 	Menyajikan aturan pengembangan pasar film China berdasarkan genre, waktu rilis, dan proporsi bahasa, yang akan sangat

			<ul style="list-style-type: none"> Analisis visual 	menguntungkan untuk praktisi film dalam menganalisis pasar, pembuatan keputusan dalam pembuatan proyek film, dan waktu rilis. Aturan ini dapat digunakan sebagai panduan di masa mendatang.
	Shaorong dkk. [6]	2018	<ul style="list-style-type: none"> Data film Big data Analisis visual 	Mengusulkan sistem analisis visual untuk big data film box-office di China
	Liao dkk. [7]	2015	<ul style="list-style-type: none"> Profil pengguna Big data 	Mendefinisikan profil <i>heavy-user</i> dan <i>non-heavy user</i> di perusahaan Chungha Telecom dari 12 aplikasi video teratas.
	Takano dkk. [8]	2017	<ul style="list-style-type: none"> Layanan streaming musik Big data Hidden Markov model Pemodelan pengguna 	Memodelkan tren musik yang didengarkan pengguna menggunakan hidden markov model. Model ini akan mengevaluasi bagaimana algoritma rekomendasi berdampak pada perilaku pengguna dalam mendengarkan musik.
Rekomendasi Konten	Zhou dkk. [9]	2016	<ul style="list-style-type: none"> Pemisahan privasi Pembelajaran online terdistribusi <i>Media cloud</i> Multimedia big data <i>Online social networks</i> (OSNs) Rekomendasi video 	Menyajikan kerangka kerja dalam merekomendasikan video pada OSNs.
	Xing dkk. [10]	2016	<ul style="list-style-type: none"> Podcast NLP Recommendation 	Mengusulkan pendekatan baru untuk memanfaatkan beragam data teks yang terasosiasi dengan item podcast untuk menyimpulkan atau memprediksi kategori konten, sehingga dihasilkan rekomendasi berbasis konten untuk item podcast yang memiliki volume besar.
	McLean dan Davis [11]	2016	<ul style="list-style-type: none"> Penanda sosial Big data Kemiripan semantik Mesin rekomendasi Model topik <i>Document embedding</i> 	Menyajikan metode baru dalam memanfaatkan metadata dengan tujuan mempelajari tautan semantik pada buku, selain itu juga menunjukkan bahwa fitur semantik yang dipelajari dapat meningkatkan rekomendasi berbasis konten.
	Shang dkk. [12]	2016	<ul style="list-style-type: none"> Video mikro Sistem rekomendasi Slope one Visualisasi data 	Memusulkan sebuah sistem rekomendasi untuk video mikro. Berdasarkan penelusuran pengguna atau riwayat menonton, sistem ini dapat merekomendasikan video favorit pada penonton. Sistem ini juga dapat mengumpulkan rekasi dan memberi saran kepada produser video dengan memberikan info berapa banyak penonton yang menyukai video mikro.
	Halder dkk. [13]	2014	<ul style="list-style-type: none"> Sistem rekomendasi Preferensi pengguna Jenis hiburan yang menarik Jenis hiburan populer 	Mengusulkan algoritma rekomendasi hiburan yang merekomendasikan jenis hiburan pada sekelompok pengguna.
Prediksi Minat	Zhu dkk [14]	2017	<ul style="list-style-type: none"> Siaran televisi Prediksi popularitas <i>Dynamic time warping</i> <i>Regresi random forest</i> <i>Gradient boosting decision tree</i> 	Menyajikan model prediksi popularitas program
Platform Media dan Hiburan	Hu [15]	2015	<ul style="list-style-type: none"> Platform <i>big data</i> Analisis TV sosial SDN (<i>Software Defined Networking</i>) <i>Crawler data</i> 	Menyajikan platform big data cloud-centric untuk analisis TV sosial, bertujuan untuk menggali persepsi sosial pada program TV dari media sosial.
Churn Pelanggan	Bi dkk [16]	2016	<ul style="list-style-type: none"> Axiomatic fuzzy sets (AFSS) mapreduce semantic-driven subtractive clustering method (SDSCM) 	Mengusulkan algoritma baru, yaitu SDSCM untuk meningkatkan akurasi pengklusteran dari SCM dan k-means. Selain itu, algoritma ini mengirangi ketidaktepatan manajemen operasi menggunakan AFS.

			<ul style="list-style-type: none"> • subtractive clustering method (SCM). 	
	Castro dan Suzuki [17]	2015	<ul style="list-style-type: none"> • Data mining • Permainan • Machine learning • Model prediktif • Wavelet transform 	Mengusulkan sebuah pendekatan untuk analisis frekuensi dari data login untuk memodelkan prediksi <i>churn</i> .

III. HASIL DAN PEMBAHASAN

Berdasarkan hasil survei penelitian-penelitian yang berfokus pada big data di industri media dan hiburan, dapat diketahui bahwa Big Data banyak memberikan keuntungan pada industri ini. Industri media dan hiburan sangat tergantung dari respon dan perilaku pengguna, sehingga memberikan kepuasan pada pengguna merupakan tujuan utama dari industri media dan hiburan.

Pemanfaatan Big Data, selain untuk memberikan keuntungan bagi pengguna, juga memberikan keuntungan bagi pihak-pihak yang terlibat pada industri media dan hiburan seperti berikut [18].

- Penerbit video: pembuat video mandiri atau pribadi, yang menerbitkan konten termasuk video, audio, teks, dan gambar.
- Pemilik media: bisnis yang memiliki hak cipta untuk menjual konten yang dapat dijual melalui ritel atau media pendistribusi konten besar-besaran.
- Perusahaan permainan: video permainan online atau offline yang dapat mencatat reaksi pemain untuk menyetel pengalaman bermain yang bagus.
- Saluran TV: saluran TV yang menyiarkan konten video yang diproduksi sendiri atau hasil beli dari pihak lain kepada masyarakat luas.

Pada penelitian survey ini, dieksplorasi beberapa cara tentang bagaimana Big Data membantu industri media dan hiburan dalam memanfaatkan banyaknya data yang mengalir dari berbagai sumber.

A. Analisis Tren

Mengetahui tren yang sedang diminati pengguna saat ini, merupakan kebutuhan penting bagi industri media dan hiburan untuk menyediakan konten dan fitur yang sesuai. Analisis tren dapat memberikan informasi mengenai perihal apa yang sedang disukai atau fitur apa yang sedang dibutuhkan masyarakat saat ini. Pada penyiaran TV, hal ini berguna untuk menaikkan nilai rating program juga dalam penargetan iklan yang efektif.

B. Rekomendasi Konten

Rekomendasi jenis hiburan adalah pekerjaan yang menantang karena terdapat perbedaan preferensi yang dirasakan pengguna satu sama lain. Misalnya, anak-anak lebih suka film animasi, remaja menyukai film petualangan, dan orang dewasa menyukai film keluarga. Selain itu preferensi hiburan dapat berubah pada periode waktu tertentu. Misalnya, mendengarkan berita atau menonton acara olahraga pada hari kerja dan menikmati film di akhir pekan atau membaca komik di masa kecil dan membaca novel saat dewasa. Preferensi hiburan juga bervariasi berdasarkan budaya. Misal, melakukan pekerjaan keagamaan atau pekerjaan sosial. Sistem rekomendasi akan sangat efektif jika dapat menghadapi tantangan sebanyak mungkin [13]. Dengan memberikan rekomendasi berkualitas tinggi kepada pengguna dapat meningkatkan keterlibatan pengguna. Maka dari itu, banyak penelitian yang fokus dalam mengusulkan teknik sistem rekomendasi dengan memanfaatkan Big Data.

C. Prediksi Konten dan Minat Pengguna

Konten media tradisional tersedia dalam format terbatas. Saat ini, konten tersebut digantikan oleh berbagai layanan media yang dibayar per tampilan, sesuai permintaan, streaming secara langsung, dan banyak lagi. Dalam proses penyampaian konten melalui format ini, penyiar juga mengumpulkan sejumlah besar data pengguna yang dapat memberikan pemahaman mendalam tentang perilaku dan preferensi.

YouTube¹ merupakan sebuah situs web berbagi video yang dibuat pada Februari 2005. Situs ini memungkinkan pengguna mengunggah, menonton, dan berbagi video. Lebih dari 1,9 Miliar pengguna

mengunjungi YouTube setiap bulannya dan setiap hari orang menonton lebih dari 1 Miliar jam video [19]. YouTube telah memiliki hasil statistik tentang penggunaannya seperti video jenis apa yang paling banyak ditonton, perangkat apa yang digunakan untuk *streaming* video, berapa lama setiap video ditonton, dan banyak lagi [18].

YouTube mengetahui begitu banyak tentang penggunaannya dengan Big Data dan analisisnya. Big Data memberikan wawasan mendalam tentang perilaku penonton YouTube dan membantunya memprediksi konten yang selaras dengan preferensi penonton [18].

D. Platform Media dan Hiburan

Saat ini, industri media dan hiburan tidak lagi dinikmati dalam format terbatas. Platform digital telah mengubah cara masyarakat dalam menikmati hiburan, mereka memberikan pengalaman yang tidak diberikan oleh saluran hiburan tradisional seperti TV dan radio.

Netflix² dan Spotify³ merupakan contoh platform hiburan yang memiliki jutaan pelanggan yang tersebar di seluruh dunia. Kedua platform ini menerapkan analisis big data untuk mendukung proses bisnis dan meraup keuntungan.

Netflix adalah layanan hiburan internet terkemuka di dunia dengan lebih dari 139 juta pelanggan berbayar di lebih dari 190 negara. Di Netflix, pelanggan dapat menikmati serial TV, film dokumenter, dan film layar lebar dengan berbagai genre dan bahasa. Pelanggan dapat menonton sebanyak yang mereka inginkan, kapan saja, di mana saja selama terhubung dengan internet. Pelanggan dapat memainkan, menghentikan sebentar, dan melanjutkan menonton, seluruhnya tanpa iklan atau komitmen [20].

Netflix menggunakan analisis big data untuk mengoptimalkan kualitas dan stabilitas aliran video dan juga untuk mengetahui preferensi hiburan dan pola menonton pelanggan. Hal ini memungkinkan Netflix untuk memberikan penawaran acara yang mungkin disukai penggunaannya. Dengan jumlah pelanggan yang besar, maka memunculkan juga data dalam jumlah besar yang dapat dianalisis untuk meningkatkan pengalaman pengguna. Netflix mengumpulkan data dari berbagai sumber untuk memprediksi jenis konten yang memiliki kemungkinan untuk menambah lebih banyak penonton hingga merekomendasikan konten kepada pengguna [20].

Spotify merupakan layanan musik *on-demand* terbesar di dunia. Perusahaan musik digital ini memiliki lebih dari 100 Juta pengguna. Dengan memiliki 10 Juta pendengar musik setiap menitnya dalam sehari, Spotify memiliki akses ke sejumlah besar informasi seperti lagu-lagu apa yang memiliki waktu putar paling banyak hingga perangkat apa yang digunakan pendengar untuk mengakses layanan. Tidak heran jika Spotify menjadi perusahaan yang digerakkan oleh data dan menggunakan data di setiap bagian organisasinya dalam membuat keputusan [21].

Sebagai contoh, salah satu fitur di Spotify yaitu Discover Weekly telah mencapai 40 Juta orang di tahun pertamanya. Setiap pengguna mendapat daftar putar yang dipersonalisasi setiap minggunya dari musik yang belum pernah mereka dengar sebelumnya di layanan, namun itu akan menjadi sesuatu yang diharapkan oleh pendengar untuk dinikmati [21].

E. Churn Pelanggan

Dengan munculnya Big Data, saat ini mungkin untuk mengetahui alasan pelanggan terus berlangganan atau berhenti berlangganan, juga mungkin untuk mengetahui program apa yang pelanggan sukai dan tidak sukai. Informasi mendalam terhadap respon mengenai harga dan model langganan juga dapat diambil dengan Big Data. Melalui analisis big data, harga konten, media konten, bahkan mode penyampaian konten dapat disesuaikan untuk mengurangi resiko *churn* pelanggan [18].

IV. KESIMPULAN

Big Data telah memberikan banyak manfaat bagi perkembangan industri media dan hiburan. Seiring dengan perubahan cara masyarakat dalam menikmati hiburan, dari saluran hiburan tradisional seperti TV dan Radio, menjadi hiburan yang dibayar per tampilan dan *streaming* secara langsung yang dapat dinikmati melalui platform-platform digital.

Pada makalah ini dilakukan survei terhadap penelitian-penelitian yang berfokus dalam menjawab tantangan dan memberikan solusi Big Data pada industri media dan hiburan. Dari hasil survei,

diketahui beberapa keuntungan-keuntungan yang didapat dari pengaplikasian analisis big data pada industri media dan hiburan adalah sebagai berikut: 1) Analisis tren; 2) Rekomendasi konten; 3) Prediksi konten dan minat pengguna; 4) Kemunculan platform media dan hiburan; dan 5) Memperkecil resiko *churn* pelanggan. Untuk penelitian lebih lanjut, dapat dilakukan survei lebih mendalam tentang cara dan teknik yang digunakan pada analisis big data.

REFERENSI

- [1] M. L. Stone, "Big Data for Media," no. November, 2014.
- [2] M. Suri and S. Narayan Singh, "The Role of Big Data in the Media and Entertainment Industry."
- [3] J. M. Cavanillas, · Edward Curry, and W. Wahlster, "New Horizons for a Data-Driven Economy A Roadmap for Usage and Exploitation of Big Data in Europe."
- [4] Oracle, "Improving Logistics & Transportation Performance with Big Data Architect ' s Guide and Reference Architecture Introduction," *Oracle White Pap.*, no. February, p. 24, 2015.
- [5] D. Wang, S. Chen, Z. Cui, H. Ding, and Z. Xie, "The Visualized Analysis System for Big Data of Movies' Box Office in China," 2019, pp. 90–94.
- [6] H. Shaorong, X. Zhifeng, and H. Jianbo, "Visual Analysis of Big Data Based on Movies Released in China," 2019, pp. 80–85.
- [7] C. H. Liao, Y. H. Lei, K. Y. Liou, J. S. Lin, and H. F. Yeh, "Using Big Data for Profiling Heavy Users in Top Video Apps," in *Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress 2015*, 2015, pp. 381–385.
- [8] M. Takano *et al.*, *Analysis of the Changes in Listening Trends of a Music Streaming Service*. 2017.
- [9] P. Zhou, Y. Zhou, D. Wu, and H. Jin, "Differentially Private Online Learning for Cloud-Based Video Recommendation with Multimedia Big Data in Social Networks," *IEEE Trans. Multimed.*, vol. 18, no. 6, pp. 1217–1229, Jun. 2016.
- [10] Z. Xing, M. Parandehgheibi, F. Xiao, N. Kulkarni, and C. Pouliot, "Content-based recommendation for podcast audio-items using natural language processing techniques," *Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016*, pp. 2378–2383, 2016.
- [11] N. McLean and J. Davis, "Utilising semantically rich big data to enhance book recommendation engines," *Proc. - 18th IEEE Int. Conf. High Perform. Comput. Commun. 14th IEEE Int. Conf. Smart City 2nd IEEE Int. Conf. Data Sci. Syst. HPCC/SmartCity/DSS 2016*, pp. 1434–1441, 2017.
- [12] S. Shang, M. Shi, W. Shang, and Z. Hong, "A micro-video recommendation system based on big data," *2016 IEEE/ACIS 15th Int. Conf. Comput. Inf. Sci. ICIS 2016 - Proc.*, pp. 1–5, 2016.
- [13] S. Halder, M. H. Seddiqui, and Y. K. Lee, "An entertainment recommendation system using the dynamics of user behavior over time," *2014 17th Int. Conf. Comput. Inf. Technol. ICCIT 2014*, no. August 2015, pp. 41–46, 2003.
- [14] C. Zhu, G. Cheng, and K. Wang, "Big Data Analytics for Program Popularity Prediction in Broadcast TV Industries," *IEEE Access*, vol. 5, pp. 24593–24601, Oct. 2017.
- [15] H. Hu, Y. Wen, Y. Gao, T. S. Chua, and X. Li, "Toward an SDN-enabled big data platform for social TV analytics," *IEEE Netw.*, vol. 29, no. 5, pp. 43–49, 2015.
- [16] W. Bi, M. Cai, M. Liu, and G. Li, "A Big Data Clustering Algorithm for Mitigating the Risk of Customer Churn," *IEEE Trans. Ind. Informatics*, vol. 12, no. 3, pp. 1270–1281, 2016.
- [17] E. G. Castro and M. S. G. Tsuzuki, "Churn Prediction in Online Games Using Players ' Login Records : A Frequency Analysis Approach," *IEEE Trans. Comput. Intell. AI Games*, vol. 7, no. 3, pp. 255–265, 2015.
- [18] S. Philips, "5 Ways Big Data Plays a Major Role in the Media and Entertainment Industry," 2017. [Online]. Available: <https://www.analyticsindiamag.com/5-ways-big-data-plays-major-role-media-entertainment-industry/>. [Accessed: 06-Apr-2019].
- [19] Youtube, "YouTube by the numbers." [Online]. Available: <https://www.youtube.com/yt/about/press/>. [Accessed: 06-Apr-2019].
- [20] Roshan, "Deciphering the unstoppable Netflix and the role of Big Data," 2018. [Online]. Available: <https://www.muvi.com/blogs/deciphering-the-unstoppable-netflix-and-the-role-of-big-data.html>. [Accessed: 06-Apr-2019].

- [21] B. Marr, "The Amazing Ways Spotify Uses Big Data, AI And Machine Learning To Drive Business Success," 2017. [Online]. Available: <https://www.forbes.com/sites/bernardmarr/2017/10/30/the-amazing-ways-spotify-uses-big-data-ai-and-machine-learning-to-drive-business-success/#36cc238b4bd2>. [Accessed: 06-Apr-2019].