

Penggabungan Teknologi Untuk Analisa Data Berbasis Data Science

Mohamad Adhisyanda Aditya, R. Dicky Mulyana, I Putu Eka, Septian Rheno Widiyanto

Program Pasca Sarjana Sistem Informasi, STMIK LIKMI Bandung, Indonesia

Email:¹adhisyanda@gmail.com,²dickymaulana@gmail.com,³ekawarmayudha@gmail.com,⁴septian.rheno@likmi.ac.id

Abstrak—Sekarang ini data berkembang menjadi life blood dalam arti sesungguhnya. Organisasi berlomba-lomba untuk dapatkan keuntungan dari informasi juga data yang dikelola. Hal tersebut menyebabkan analisa data menjadi suatu hal yang penting bagi sebuah organisasi. Seiring dengan perkembangan data yang sangat pesat, maka metode yang digunakan untuk memanage data pun tumbuh dengan pesat pula. Pada era saat ini, manusia dihadapkan dengan jumlah besar dari data yang akan sangat sulit untuk dikelola dan di analisa dengan metode yang digunakan sebelumnya. Berkembangnya teknologi juga ikut menjadi pemacu untuk dilakukan analisis dengan metode baru. Sekarang ini kita tidak saja hanya berhadapan dengan data yang bersifat terstruktur, tapi juga dengan data yang tidak terstruktur.

Kata Kunci: Data Science, Machine Learning, Deep Learning, Big Data

1. PENDAHULUAN

Para pelaku bisnis mengartikan Revolusi 4.0 yaitu dengan mengambil keputusan secara cepat dan tepat. Data merupakan penunjang pengambilan keputusan secara cepat. Atau dikenal dengan istilah Data Driven Decision Making (DDDM). Permasalahannya adalah diperlukan waktu dan tenaga yang tidak sedikit dalam menganalisis data sehingga keputusan yang bersifat mendesak terkadang mengandalkan intuisi.

Kemajuan teknologi internet, media digital seperti audio, video, dan teks dikirimkan melalui Internet dengan lebih mudah. Saat memiliki data, pertanyaan selanjutnya yang ada adalah, apa yang dapat dimanfaatkan dengan data ini? Jika hanya menyimpan data tersebut secara historical, maka yang akan didapatkan hanyalah storage yang semakin lama akan semakin berkurang. Dalam sebuah artikel di media forbes menyebutkan bahwa 90% data merupakan data yang di generate dalam kurun waktu 2 tahun terakhir. Hal tersebut dapat menggambarkan bahwa pertumbuhan data saat ini sangatlah pesat.

Beberapa fakta yang terjadi telah dipaparkan, maka dapat disimpulkan bahwa saat ini organisasi yang dapat memanfaatkan data yang dimilikinya dalam waktu yang tepat dengan metode yang tepat akan sangat mendapatkan keuntungan dari pengelolaan data tersebut

Salah satu pengertian tentang data science adalah suatu disiplin ilmu yang khusus mempelajari data, khususnya data kuantitatif (data numeric), baik yang terstruktur ataupun yang tidak terstruktur. Berbagai subjek dalam data science meliputi semua proses data, mulai dari pengumpulan data, analisis data, pengolahan data, manajemen data, kearsipan, pengelompokan data, penyajian data, distribusi data, dan cara mengubah data menjadi kesatuan informasi yang dapat

dipahami semua orang. Ilmu-ilmu yang menjadi penunjang utama dalam ilmu data terdiri dari matematika, statistika, ilmu komputer, sistem informasi, manajemen, ilmu informasi, termasuk juga ilmu komunikasi dan kepustakaan. Bahkan ilmu ekonomi, terutama ilmu bisnis, juga berperan penting dalam ilmu data.

Orang yang berkerja dengan Data Science, disebut data scientist. Data Scientist itu sudah ada yang sering disebut dengan statistikawan. Oleh karena itu, tidak mengherankan jika data scientist sekarang lebih sering menciptakan algoritma-algoritma di dalam program komputer agar data yang masuk dapat langsung diolah sendiri oleh komputer tersebut.

Data Science merupakan ilmu yang sangat cepat berkembang dan di kembangkan oleh organisasi yang menggunakan konsep DDDDM. Beberapa komponen teknologi yang sangat menunjang data science diantaranya adalah big data, machine learning, dan internet of things. Saat semua komponen tersebut dapat di kombinasikan dan di olah untuk kebutuhan organisasi, maka ke “sexy” an data untuk organisasi pun dapat terasa.

2. METODOLOGI PENELITIAN

Data Science memiliki siklus hidup atau stage dalam implementasinya. Siklus hidup yang dijalani dalam data science terdiri dari 5 stage yaitu:

- a. Capture
Siklus ini berhubungan dengan pengumpulan data
- b. Maintain
Siklus ini berhubungan dengan data warehouse, data staging, data cleansing, arsitektur data.
- c. Process
Siklus ini berhubungan dengan data pemrosesan data, data mining, data modeling, data summarizing.
- d. Analyze
Siklus ini berhubungan dengan analisa data, data predictive.
- e. Communicate
Siklus ini berhubungan dengan visualisasi data, data reporting, pengambilan keputusan.



Gambar 1. Siklus Hidup Data Science

3. ANALISA DAN PEMBAHASAN

Seperti yang telah di jelaskan sebelumnya, bahwa dalam implementasi data science, ada sejumlah teknologi yang sangat menunjang, seperti big data, machine learning, dan internet of things.

a. Internet Of Things

Internet of Things atau dikenal juga dengan singkatan IoT, merupakan sebuah konsep yang bertujuan untuk memperluas manfaat dari konektivitas internet yang tersambung secara terus-menerus. Adapun kemampuan seperti berbagi data, remote control, dan sebagainya, termasuk juga pada benda di dunia nyata. Contohnya bahan pangan, elektronik, koleksi, peralatan apa saja, termasuk benda hidup yang semuanya tersambung ke jaringan lokal dan global melalui sensor yang tertanam dan selalu aktif.

Berdasarkan prediksi perusahaan IT terkenal Cisco, pada tahun 2020, akan ada 50 milyar alat-alat yang terhubung dengan Internet, termasuk 400 juta alat-alat tersebut adalah perangkat dalam kategori wearables. Beberapa contoh konkrit dari “wearable” yang mulai dipasarkan di dunia adalah : Google Glass, Google Nest, Nike Fit, dan Samsung Smart Watch. Tidak hanya wearables, Samsung juga mulai merambah dan mengembangkan teknologi IoT di bidang consumer appliances seperti: Smart Air Conditioner, Smart TV, Smart Refrigerator

Dengan semakin berkembangnya infrastruktur internet, maka menuju babak berikutnya, bukan hanya smartphone atau komputer saja yang dapat terkoneksi dengan internet. Namun berbagai macam benda nyata akan terkoneksi dengan internet. Sebagai contohnya dapat berupa: mesin produksi, mobil, peralatan elektronik, peralatan yang dapat dikenakan manusia (wearables), dan termasuk benda nyata apa saja yang semuanya tersambung ke jaringan lokal dan global menggunakan sensor dan atau aktuator yang tertanam. Produk dibangun dengan kemampuan komunikasi M2M yang sering disebut dengan sistem cerdas atau “smart”. (contoh: smart label, smart meter, smart grid sensor).

Kata “internet” dalam Internet of Things bermakna jaringan. Perangkat IoT tidak perlu harus terhubung ke internet, tetapi perlu terhubung ke sebuah jaringan. (Boleh jadi ini jaringan tertutup.) Ada beberapa protokol yang dapat menghubungkan perangkat ke jaringan. Protokol yang digunakan adalah TCP/IP. Lebih spesifik lagi, untuk jaringan nirkabel (wireless), protokol yang banyak digunakan adalah WiFi (IEEE 802.4). Protokol Bluetooth (lebih spesifik lagi BLE) mulai banyak digunakan pula pada beberapa aplikasi yang lingkup fisiknya lebih dekat. Untuk aplikasi yang menghubungkan perangkat ke jaringan yang lebih jauh, protokol LoRa sedang naik daun.

Salah satu komponen utama dari IoT adalah development boards yang digunakan sebagai basis. Kedalam board inilah program atau yang lebih dikenal dengan istilah sketch akan di masukkan. Dengan program tersebutlah board akan di control untuk melakukan berbagai macam aktifitas yang diinginkan.

Arduino merupakan Board yang digunakan. Kesuksesan dari Arduino adalah keterbukaan desain dari board (hardware) dan software (dalam bentuk Arduino IDE). Akibatnya banyak orang dapat memproduksi board Arduino dengan harga yang terjangkau.

Board yang kemudian terkenal adalah board yang berbasis ESP8266. Kelebihan dari board berbasis ESP8266 adalah board tersebut sudah memiliki modul WiFi (802.11 b/g/n, dengan WPA/WPA2), yang akan memudahkan dalam menghubungkan perangkat ini ke jaringan. (Pembahasan mengenai aspek jaringan yang lebih rinci ada pada bagian terpisah.) Ada banyak implementasi dari board berbasis ESP8266 ini, antara lain dari Espressif, NodeMCU. Kepopuleran dari board berbasis ESP8266 ini membuat harganya menjadi murah.

Aspek harga sangat menentukan dalam pengembangan sebuah sistem. Untuk sistem dengan jumlah titik (node) yang sedikit, perbedaan harga tidaklah terlalu signifikan. Untuk sistem dengan jumlah titik (node) yang banyak, maka aspek harga sangat menentukan. Sebagai contoh untuk sistem dengan 1000 titik , dimana masing-masing titik membutuhkan setidaknya satu board, perbedaan harga Rp. 10.000,- akan menjadikan beda sepuluh juta rupiah. Padahal ada board yang lebih bagus tetapi harganya berbeda cukup signifikan. Sebagai contoh, ada salah satu board (Puck.js) yang berharga Rp. 300.000,- untuk satu unitnya.

Sketch merupakan sebutan untuk program atau kumpulan kode yang digunakan untuk mengontrol development board. Sketch dapat ditulis menggunakan berbagai IDE atau text editor, contohnya adalah ArduinoIDE, Visual Studio Code(dengan bantuan platformIO plugin). IDE yang banyak digunakan salah satunya adalah ArduinoIDE.

Saat ArduinoIDE pertama kali dijalankan, maka default struktur yang akan di tampilkan adalah seperti gambar di atas. Terdapat dua fungsi awal yang kemudian akan diisi oleh program. Dua fungsi awal tersebut adalah void setup dan void loop. Void setup merupakan fungsi yang hanya akan di eksekusi 1x di awal. Sedangkan void loop merupakan fungsi yang akan terus menerus di eksekusi selama board dijalankan.

ArduinoIDE juga dapat digunakan untuk mengupload sketch ke beberapa development board selain Arduino, contohnya adalah wemos. Untuk development board selain Arduino, perlu dilakukan beberapa konfigurasi untuk menyesuaikan environment pada IDE dengan ArduinoIDE.

Salah satu media untuk menghubungkan peralatan IoT dengan lingkungannya adalah sensor. Sensor dapat bekerja untuk merubah sesuatu yang di tangkap oleh sensor tersebut menjadi bentuk data yang kemudian dapat di proses. Contohnya adalah sensor temperature yang dapat menangkap suhu, lalu merubahnya menjadi data suhu yang kemudian dapat di proses.

Protocol-protocol yang digunakan dalam IoT, kebanyakan digunakan untuk media komunikasi peralatan IoT dengan dunia luar seperti pengguna, dan server penyimpanan data yang di hasilkan dari sensor.

Terdapat beberapa protokol yang dapat digunakan dalam pengaplikasian IoT, Contohnya yang paling sering di gunakan adalah protocol Wifi (IEEE 802.11). Protocol-protocol tersebut digunakan sebagai metode komunikasi antar perangkat yang digunakan dalam IoT. Contohnya, saat membuat sebuah sensor cuaca, diperlukan sebuah server untuk menampung nilai yang dihasilkan dari sensor tersebut untuk kemudian data yang di kumpulkan sensor tersebut dapat di oleh lebih lanjut. Komunikasi yang digunakan untuk menghubungkan sensor dan development board ke jaringan (baik itu internet atau local) dapat menggunakan protocol Wifi.

Protocol lain yang banyak di gunakan dalam impelentasi adalah MQTT (Message Queue Telemetry Transport). MQTT adalah protocol publisher dan subscriber lightweight. Protocol ini sangat berguna untuk digunakan dalam device yang memiliki power yang rendah tetapi dapat di aplikasikan dalam banyak scenario. Protocol ini pada dasarnya adalah publish dan subscribe sebuah pesan kedalam topik. Beberapa client terhubung kesebuah broker dan mengambil sebuah pesan berdasarkan topiknya. Client yang lainnya dapat juga terhubung ke broker sebagai subscriber dan mengirimkan pesan berdasarkan topiknya.

Dalam siklus hidup data science, IoT dapat berperan dalam siklus capture. Dengan memanfaatkan development board dan sensor yang dihubungkan, maka IoT device dapat mengumpulkan data dan informasi dengan waktu yang cepat dan frekuensi yang juga cepat.

b. Big Data

Definisi terminologi Big Data secara konsep dapat dikatakan belum memiliki acuan ilmiah yang baku. Diebold menegaskan bahwa apa yang disebut Big Data sekarang berbeda dengan definisi Big Data pada 15 tahun yang lalu. Beberapa rujukan tentang Big data pada awal atau sebelum tahun 2000 sangatlah menarik namun belum menyakinkan. Kesadaran akan kebutuhan pengelolaan dan analisis data berukuran besar telah merebak di lingkungan Silicon Graphics (SGI) pada pertengahan tahun 1990.

Press bahkan menyebutkan jauh sebelum itu, kesadaran tersebut sudah ada. Gandomi dan Haider yang juga merujuk ke Diebold, menambahkan bahwa terminologi ini mulai banyak diperbincangan pada naskah publikasi ilmiah pada tahun 2012 dan 2013 sebanyak 2-3 kali lipat dari jumlah publikasi tahun 2011 (berdasar dokumen yang mengandung big data dalam ProQuest Research Library).

Pada umumnya, penggunaan parameter ukuran data lebih umum digunakan untuk menjawab apakah big data atau tidak. Parameter ukuran (volume) data bukanlah satu-satunya acuan. Gandomi and Haider merujuk Laney pada penggunaan parameter: Three-V, yaitu Volume, Variety, dan Velocity (Gartner menyebutnya sebagai bagian – Three-Parts). Dari sekian banyak definisi terkait Big Data, beberapa di antaranya menggunakan Three-V sebagai penekannya.

Volume merujuk pada ukuran besarnya data yang dikelola (dalam satuan MB, GB, TB, PB, ZB). Batasan ukuran yang dapat disebut Big Data masih dapat dikatakan beragam. Furht and Villanustre, juga Driscoll menyatakan ukuran minimum sebuah sistem Big Data adalah dari Terabytes (TB) sampai Petabyte (PB). Variety merujuk pada tingkat keragaman struktur dalam dataset, dapat berupa terstruktur (contoh tabel), semi terstruktur (dokumen XML), dan tidak terstruktur (dokumen, email, text message, audio, video, gambar, graphic, dan lainnya). Velocity merujuk pada cara dan tingkat kecepatan penerimaan data, termasuk apakah melalui proses batch, near time, real time, dan stream.

Furht and Villanustre, pada Tabel I memberikan ilustrasi perbedaan antara data tradisional dengan big data. Gambaran fakta terkait ketiga dimensi tersebut, antara lain: Facebook dilaporkan mengelola: 2.5 Miliar item konten yang dibagikan, 2.7 Miliar Likes, 300 juta foto di upload, 500+ TB data diproses, 70 ribu eksekusi query, 100+ PB ruang penyimpanan dalam sebuah HDFS cluster, dan 105 TB data tertelusuri melalui Hive. Data serupa juga ditunjukkan oleh Twitter tahun 2010: 1+ miliar tweet dalam seminggu, rata-rata 50 juta tweet terkirim per hari.

Tabel 1. Perbandingan Tradisional dan Big Data

	Tradisional Data	Big Data
Volume	GB	TB dan PB
Laju	Data Per jam, per hari	lebih cepat
Struktur Data	Terstruktur	semi atau tidak terstruktur
Sumber Data	Terpusat	Tersebar
Intergrasi Data	Mudah	Sulit
Penyimpan Data	RDBMS	HDFS, NoSQL
Akses Data	Interaktif	real time

Selain ketiga dimensi tersebut, beberapa pemain industry menyebutkan dimensi lainnya, yaitu: Veracity (ketelitian) disebutkan oleh IBM dan lebih merujuk terhadap kualitas data; Variability (dan complexity) dinyatakan oleh SAS untuk menunjukkan variasi kecepatan aliran data yang berasal dari beberapa sumber; dan Value dinyatakan oleh Oracle untuk menekankan bahwa data mentah dari sumber lebih cenderung tidak berkualitas (low value density) sehingga perlu dianalisis untuk mendapatkan high value.

c. Machine Learning

Beberapa tahun terakhir, banyak yang mulai mempelajari Machine Learning. Ada beberapa pertanyaan yang sering disampaikan:

1. Apa itu Machine Learning dan bagaimana bedanya dengan Big Data dan Business Analytics?
2. Apa perbedaan antara Machine Learning, Data Analysis, Data Mining, Data Science dan Artificial Intelligence?

Secara definisi, machine learning atau pembelajaran mesin adalah ilmu atau studi yang mempelajari tentang algoritma dan model statistik yang digunakan oleh sistem komputer untuk melakukan task tertentu tanpa instruksi eksplisit.

AI ini mengacu pada prosedur pemrograman komputer (machine) untuk mengambil suatu yang rasional. Misalnya, mesin dapat menimbulkan alarm jika parameter mengatakan 'X' melintasi ambang batas tertentu yang pada gilirannya dapat mempengaruhi hasil proses terkait.

Machine Learning adalah subset dari AI dimana mesin dilatih untuk belajar dari pengalaman masa lalu. Pengalaman masa lalu dikembangkan melalui data yang dikumpulkan, kemudian menggabungkan dengan algoritma (seperti Naïve

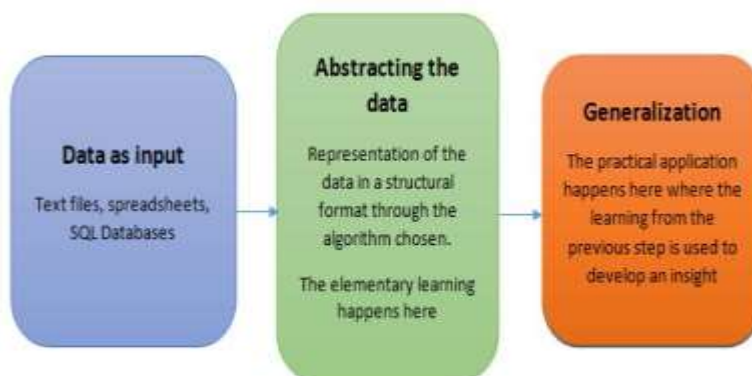
Bayes, Support Vector Machine (SVM)) untuk memberi hasil akhir.

Untuk mengidentifikasi email spam, pengguna dapat menggunakan algoritma Machine Learning yang dikenal sebagai Naïve Bayes yang akan memeriksa frekuensi kiriman spam masa lalu. Untuk mengidentifikasi email baru sebagai spam, Naïve Bayes menggunakan teori statistik Baye's Theorem (umumnya dikenal sebagai probabilitas bersyarat). Oleh karena itu, dapat dikatakan algoritma Machine Learning menggunakan konsep statistik untuk melakukan pembelajaran mesin.

Deep Learning dikaitkan dengan algoritma jaringan saraf tiruan – Artificial Neural Network (ANN) yang menggunakan konsep otak manusia untuk memudahkan pemodelan fungsi yang berubah-ubah. ANN membutuhkan sejumlah besar data dan algoritma ini sangat fleksibel dalam hal menghasilkan keluaran secara bersamaan.

Data Mining digunakan untuk mencari informasi yang spesifik, sedangkan Machine Learning berkonsentrasi untuk melakukan tugas tertentu. Sebagai contoh untuk membantu perbedaan antara Machine Learning dan Data Mining, mengajarkan seorang cara menari adalah Machine Learning, sedangkan menggunakan seseorang untuk mencari pusat tarian terbaik di kota adalah Data Mining.

Machine Learning melibatkan proses struktural, setiap tahap membangun versi mesin yang lebih baik. Untuk penyederhanaan, proses Machine Learning dapat dibagi menjadi 3 bagian:



Gambar 2. Tiga Bagian Machine Learning

Jenis Algoritma Machine Learning yang dapat digunakan diantaranya :

1. Model Supervised Learning / Predictive

Model ini digunakan untuk memprediksi hasil masa depan berdasarkan data historis. Model prediktif biasanya diberi instruksi yang jelas sejak awal seperti apa yang perlu dipelajari dan bagaimana itu perlu dipelajari. Algoritma pembelajaran ini disebut Supervised Learning. Sebagai contoh: Supervised Learning digunakan saat perusahaan pemasaran mencoba untuk mengetahui pelanggan mana yang cenderung berpindah atau mencari

supplier lain. Algoritma ini juga dapat digunakan untuk memprediksi kemungkinan terjadinya bahaya seperti gempa bumi, tornado dan lain-lain, dengan tujuan untuk mengetahui Total Nilai Asuransi. Beberapa contoh algoritma yang digunakan adalah: Nearest Neighbour, Naïve Bayes, Decision Tree, Regression, dan lain-lain.

2. Model UnSupervised Learning/Descriptive

Model ini digunakan untuk melatih tidak memiliki target yang ditetapkan dan tidak ada faktor yang penting dari yang lainnya. Sebagai contoh penggunaan model unsupervised learning ini, bila seorang penjual pengecer ingin mengetahui kombinasi produk apa yang cenderung lebih sering dibeli konsumen. Di industri farmasi, digunakan untuk memprediksi penyakit mana yang mungkin terjadi bersamaan dengan diabetes. Contoh algoritma yang digunakan di model ini: K-Means Clustering Algorithm.

3. Reinforcement Learning (RL)

Model ini adalah contoh pembelajaran mesin dimana mesin dilatih untuk mengambil keputusan spesifik berdasarkan kebutuhan bisnis dengan tujuan utama untuk memaksimalkan efisiensi (kinerja). Ide dari Reinforcement learning ini adalah mesin/perangkat lunak melatih dirinya secara terus menerus berdasarkan lingkungan yang dipengaruhinya, dan menerapkan pengetahuan yang diperoleh untuk memecahkan masalah bisnis. Proses belajar yang terus-menerus ini memastikan lebih sedikit keterlibatan manusia sehingga akan banyak menghemat waktu. Contoh algoritma yang digunakan dalam RL adalah Markov Decision Process.

Untuk membedakan antara Supervised Learning dan Reinforcement Learning, dapat dicontohkan, sebuah mobil menggunakan Reinforcement learning untuk membuat keputusan rute mana yang harus ditempuh, kecepatan berapa yang harus dikemudikan, beberapa pertanyaan tersebut diputuskan setelah berinteraksi dengan lingkungan. Sedangkan memperkirakan ongkos taksi dari satu tempat ke tempat lain adalah Supervised Learning.

Google dan Facebook adalah dua contoh perusahaan yang menggunakan Machine Learning secara ekstensif untuk mendorong iklan masing-masing ke pengguna yang relevan. Contoh penggunaan Machine Learning yang lainnya adalah:

4. Layanan Perbankan & Keuangan

Machine Learning dapat digunakan untuk memprediksi pelanggan yang cenderung gagal membayar pinjaman atau tagihan kartu kredit. Ini sangat penting karena Machine Learning akan membantu bank untuk mengidentifikasi nasabah yang dapat diberikan pinjaman dan kartu kredit.

5. Kesehatan

Digunakan untuk mendiagnosis penyakit mematikan (misalnya kanker) berdasarkan gejala pasien dan menghitungnya dengan data terakhir dari jenis pasien yang sama.

6. Ritel

Digunakan untuk mengidentifikasi produk yang lebih sering dijual (bergerak cepat) dan produk yang lambat. Hal ini membantu memutuskan jenis produk yang akan ditampilkan atau dikeluarkan dari rak. Selain itu, algoritma Machine Learning dapat digunakan untuk menemukan dua atau lebih produk yang dijual bersama. Hal ini dilakukan untuk merangsang inisiatif loyalitas pelanggan yang pada gilirannya membantu para peritel untuk mengembangkan pelanggan setia.

Dalam siklus hidup data science, machine learning sangat berperan dalam siklus process and analyze.

4. KESIMPULAN

Dengan berkembangnya data yang sangat pesat di abad ke-21 ini, menjadikan data science sebagai metode yang sangat membantu dalam pengelolaan data. Tidak heran yang menyebutkan bahwa data scientists merupakan pekerjaan yang sangat “sexy” di abad 21. Dengan data science dan teknologi yang mendukungnya, terutama data yang berukuran besar dan data yang tidak terstruktur dapat diolah sehingga menghasilkan analisa yang melebihi cara analisa dan pengolahan data sebelumnya, seperti data warehouse dan data mining.

REFERENCES

- [1] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," International Journal of Information Management, vol. 35, pp. 137-144, 2015.
- [2] Berkeley School of Information, "What is Data Science". [Online]. Available: <https://datascience.berkeley.edu/about/what-is-data-science/>
- [3] B. Furht and F. Villanustre, "Introduction to Big Data," in Big Data Technologies and Application, Switzerland, Springer International Publishing, 2016, pp. 3-11.
- [4] B. Marr, "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read". Forbes, 21 May 2018. [Online]. Available: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#45d56c9160ba>
- [5] B. Rahardjo, "Desain dan Implementasi Aplikasi IOT Untuk Sensor Cuaca"[Online]. Available: <https://github.com/rahard/makalah-iot-weather/blob/master/makalah-iot.pdf>
- [6] C. Chan, "What Facebook Deals with Everyday: 2.7 Billion Likes, 300 Million Photos Uploaded and 500 Terabytes of Data," Gizmodo, 22 Agustus 2012. [Online]. Available: <https://gizmodo.com/5937143/what-facebook-deals-with-everyday-27-billion-likes-300-million-photos-uploaded-and-500-terabytes-of-data>
- [7] C. Chan, "What Facebook Deals with Everyday: 2.7 Billion Likes, 300 Million Photos Uploaded and 500 Terabytes of Data," Gizmodo, 22 Agustus 2012. [Online]. Available: <https://gizmodo.com/5937143/what-facebook-deals-with-everyday-27-billion-likes-300-million-photos-uploaded-and-500-terabytes-of-data>

- [8] C. Chan, "What Facebook Deals with Everyday: 2.7 Billion Likes, 300 Million Photos Uploaded and 500 Terabytes of Data," Gizmodo, 22 Agustus 2012. [Online]. Available: <https://gizmodo.com/5937143/what-facebook-deals-witheveryday-27-billion-likes-300-million-photos-uploaded-and500-terabytes-of-data>.
- [9] D. Laney, "3-D data management: Controlling data volume, velocity and variety. Application Delivery Strategies by META Group Inc.," 6 Februari 2001. [Online]. Available: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3DData-Management-Controlling-Data-Volume-Velocity-andVariety.pdf>.
- [10] F. X. Diebold, "A Personal Perspective on the Origin(s) and Development of "Big Data": The Phenomenon, the Term, and the Discipline," PIER Working Paper No. 13-003, 26 November 2012. [Online]. Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2202843.
- [11] Gartner Inc., "Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three "V"s," Forbes, 27 Maret 2013. [Online]. Available: <https://www.forbes.com/sites/gartnergroup/2013/03/27/gartner-s-big-data-definition-consists-of-three-parts-not-to-be-8confused-with-three-vs/#2abd0a8042f6>
- [12] G. Press, "12 Big Data Definitions: What's Yours?," Forbes, September 2014. [Online]. Available: <https://www.forbes.com/sites/gilpress/2014/09/03/12-big-datadefinitions-whats-yours/#519a377313ae>.
- [13] G. Press, "A Very Short History Of Big Data," Forbess, 9 Mei 2013. [Online] Available: <https://www.forbes.com/sites/gilpress/2013/05/09/a-very-shorthistory-of-big-data/>.
- [14] IBM Big Data & Analytics Hub, "The Four V's of Big Data," IBM, [Online]. Available: <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>.
- [15] M. E. Driscoll, "How much data is "Big Data"? Is there classification for various levels of "Big Data" by amount of data processed or other constraints, like for example throughput? What's the minimum data size which still qualifies as a "Big Data"?," Quora, 25 Desember 2010. [Online]. Available: <https://www.quora.com/How-much-data-is-BigData-Is-there-classification-for-various-levels-of-Big-Data-byamount-of-data-processed-or-other-constraints-like-forexample-throughput-What%E2%80%99s-the-minimum-datasize-which-still-qualifies-as-a-Big-Data%E2>.
- [16] M. L. Brodie, "What is Data Science?," researchgate publication [Online]. Available: https://www.researchgate.net/publication/333752364_What_Is_Data_Science
- [17] Oracle, "Bringing the Value of Big Data to the Enterprise," [Online]. Available: <http://www.oracle.com/us/products/database/big-dataappliance/value-of-big-data-brief-2008771.pdf>.
- [18] P. R. Choudhury, "Machine Learning Basics for a newbie". [Online]. Available: <https://www.analyticsvidhya.com/blog/2015/06/machine-learning-basics/>
- [19] SAS Institute Inc., "Big Data: What it is and why it matters," SAS Institute Inc., [Online]. Available: https://www.sas.com/en_id/insights/big-data/what-is-bigdata.html.
- [20] S. Budi, "Pembelajaran Big Data" [Online]. Available: https://www.researchgate.net/publication/329029812_Pembelajaran_Big_Data
- [21] Widiyanto, Septian Rheno. (2017). Algoritma Steganografi dengan Metode Spread Spectrum Berbasis PCMK. Jurnal Multinetics. Vol 3. No.2. <https://doi.org/10.32722/multinetics.Vol3.No.2.2017.pp.32-37>
- [22] Widiyanto, Septian Rheno. (2015). Perancangan Jaringan WLAN di PT. Gemopia Jewellery Indonesia. Jurnal Multinetics. Vol.1, No. 2. <https://doi.org/10.32722/multinetics.Vol1.No.2.2015.pp.50-53>
- [23] Widiyanto, Septian Rheno. (2020). Algoritma B217AN menggunakan Metode Spread Spectrum Berbasis PCMK/PCMB. Seminar Nasional Teknik Elektro, Prosiding SNTE Vol 5, No. 2.
- [24] Widiyanto, Septian Rheno. Desain Algoritma Steganografi dengan Metode Spread Spectrum Berbasis PCMK (Permutasi Chaotic Multiptaran Mengecil dan Membesar) Yang Tahan Terhadap Gangguan. Prodi Teknologi Rekayasa Perangkat Lunak Politeknik Enjinereng Indorama Kembang Kuning Ubrug Jatiluhur, Purwakarta. p- ISSN : 2407 – 184 e ISSN : 2460 – 8416, 2018.
- [25] Widiyanto, Septian Rheno. (2018). Desain dan Analisa Algoritma Steganografi dengan Metode Spread Spectrum Berbasis PCMK (Permutasi Chaotic Multiptaran Mengecil dan Membesar) Menggunakan Matlab. Jurnal Elektra. Vol. 3 No. 1. ISSN:2503-0221.