

PERBEDAAN DATA ENGINEER, DATA SCIENTIST DAN DATA ANALYST

Ismail Setiawan

Fakultas Sains dan Teknologi, Universitas Aisyiyah Surakarta

email: ismail@aiska-university.ac.id

Abstrak

Seseorang yang ahli dalam keterampilan analisis data hanyalah keterampilan dasar seorang insinyur data. Keahlian statistik digunakan untuk memproses data baca dan tag, serta untuk mengkategorikan data. Karena erat kaitannya dengan pemodelan yang dibuat untuk menguji algoritma pada level data scientist. Model yang dibuat pada fase data scientist digunakan sebagai alat dalam fase business intelligence. Pada tahap akhir ini, eksekusi yang akan dilakukan harus memberikan dampak positif dan keuntungan yang besar bagi sebuah instansi.

Kata kunci : ilmuwan data, statistik, analisis data, intelijen bisnis, insinyur data

Abstract

A person skilled in data analysis skills is just the basic skill of a data engineer. Statistical skills used to process read and tag data, and to categorize data. Because it is closely related to the modeling made to test algorithms at the data scientist level. The model created in the data scientist phase is used as a tool in the business intelligence phase. At this final stage, the execution that will be carried out must have a positive impact and big profits for an agency.

Keywords: data scientist, statistic, data analyst, business intelligent, data engineer .

PENDAHULUAN

Di era digital, profesi yang sering dibicarakan adalah data scientist, data analyst, dan data engineer. Ketiga profesi ini erat kaitannya dengan perkembangan teknologi dan pengolahan data. Gaji untuk data scientist, data analyst (Kelleher & Tierney, 2018), dan data engineer (Dunn & Davis, 2017) juga tidak main-main dan bisa berkisar dari puluhan hingga ratusan juta rupiah per bulan.

Seorang data scientist bertanggung jawab untuk membersihkan, memproses, dan memproses data besar yang dikumpulkan oleh para data engineer di sebuah perusahaan (Dinov, 2018). Data scientist juga sering perlu melakukan eksperimen untuk menguji dan memberikan saran yang paling tepat untuk pengembangan suatu organisasi, perusahaan, dan unit bisnis.

Data scientist sering dihadapkan pada pertanyaan seperti Berapa banyak jenis pengguna yang dimiliki perusahaan Dan

Bisakah membuat model yang memprediksi produk yang akan laku jika dijual ke target pasar tertentu. Intinya, tugas seorang data scientist adalah menarik kesimpulan berdasarkan data dalam jumlah besar yang dapat dicerna dan diterima oleh semua orang (Swarniti, 2021).

Data scientist dihadapkan dengan program pemrosesan data seperti SQL dan Python setiap hari. Mereka yang akan terjun ke dunia data scientist harus mahir setidaknya dalam bidang pemrograman data, komunikasi, matematika, statistik, dan ekonomi (Cao, 2017). Kebutuhan dunia industri terhadap seorang data engineer saat ini tidak dapat dipungkiri (Swarniti, 2019). Keberadaan seorang data engineer mampu menyelamatkan perusahaan dari kebijakan yang salah terhadap mitra kerjanya. Misalkan supermarket yang akan memberikan promosi kepada pelanggannya (Swarniti, 2021).

MASALAH, TARGET DAN LUARAN

Permasalahannya adalah apakah seorang yang mahir dalam data scientist dapat dikatakan sebagai data engineer. Masih ada beberapa tahap lagi untuk menjadi data engineer. Mereka yang ingin dikatakan sebagai data engineer harus bisa menjadi data analis.

Pekerjaan data analyst mengharuskan mereka menangani sejumlah besar data untuk membersihkan, menganalisis, dan memvisualisasikannya. Tugas data analyst adalah memperoleh informasi tentang perkembangan bisnis di masa depan dari berbagai aspek, yang kemudian diteruskan ke data engineer. Pekerjaan data analyst juga bertanggung jawab untuk memproses materi yang disediakan untuk membuat eksperimen dan menentukan strategi bisnis tambahan (Santika, 2020). Seorang data analyst mungkin akan menghabiskan hari melakukan visualisasi data yang menghubungkan tim pemasaran, tim penjualan, tim teknis, dan strategi bisnis.

Data analyst juga bertanggung jawab untuk menjawab pertanyaan seperti Bagaimana cara menjelaskan kepada manajemen bahwa kenaikan biaya mempengaruhi jumlah pelanggan dan Apa yang mendorong pertumbuhan bisnis perusahaan (Alonso-Fernandez et al., 2019). Untuk melaksanakan pekerjaan sehari-hari mereka, data analyst menggunakan program Excel, Tableau dan SQL. Data analyst harus menguasai istilah teknis dan alat untuk membuat grafik / infografis.

METODE PELAKSANAAN

Seorang data engineer adalah mereka yang mampu mengembangkan dan membuat desain arsitektur manajemen data dan memonitor infrastrukturnya di dalam sebuah perusahaan (Santika, 2021). Kelola jalur data untuk perusahaan yang menangani data dalam jumlah besar dan juga harus memastikan bahwa data dapat dikumpulkan dan diakses secara efisien dari sumbernya, dibersihkan, dan diproses seperlunya adalah sebagian

kemampuan dari seorang data engineer (Dunn & Davis, 2017).

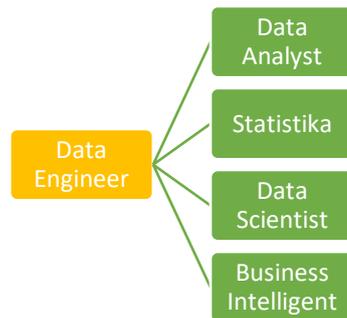
Tujuannya adalah untuk menciptakan dan mengoptimalkan sistem bisnis yang memungkinkan ilmuwan dan data analyst dalam melakukan pekerjaannya. Seorang data engineer juga harus memiliki pengetahuan tentang pemrograman, data besar, dan matematika. Sebagai seorang data engineer, menggunakan program seperti NoSQL, Hadoop, dan Python adalah hal mendasar. Seorang data engineer juga harus mahir dalam database, SQL, alat ETL, skrip shell, dan pemrograman dasar.

Meskipun ada beberapa perbedaan antara data engineer dan data scientist dan data analis, ketiga profesi tersebut masih terkait dan terkait satu sama lain. Data analyst dan data scientist tidak akan dapat bekerja tanpa data engineer. Sementara itu, data engineer tidak akan berfungsi maksimal tanpa data analyst dan data scientist.

Saat ini banyak sekali lowongan yang dibuka untuk ketiga profesi tersebut. Selain itu, ada begitu banyak bisnis yang membutuhkan seperti perbankan, kesehatan, pendidikan, media, perjalanan dan transportasi dan logistik. Oleh karena itu, seorang data scientist memiliki peluang besar untuk menjadi data scientist atau data analyst.

HASIL PEMBAHASAN

Masih ada 2 hal penting agar seseorang dapat dikatakan sebagai data engineer (Roy et al., 2017). Selain kemampuan akan data analyst dan data scientist, kemampuan lain yang harus dimiliki adalah statistika (Hao & Ho, 2019) dan business intelligent (Arnott et al., 2017). Secara urutan kemampuan dasar seorang data engineer hingga kemampuan akhir yang harus dimiliki dapat dilihat pada gambar 1.



Gambar 1. Urutan kemampuan untuk menjadi data engineer

KESIMPULAN DAN SARAN

Seseorang yang ahli dalam kemampuan data analyst hanyalah kemampuan dasar dari seorang data engineer. Keahlian dalam statistika digunakan mengolah membaca dan melabeli dan mengkategorikan data (Santika, 2017). Karena hal tersebut sangat berhubungan dengan pemodelan yang akan dibangun untuk menguji algoritma pada tahap data scientist. Model yang terbangung pada tahap data scientist akan di jadikan sebagai alat bantu dalam tahap business intelligent (Richards et al., 2019). Pada tahap akhir ini lah eksekusi yang akan dilakukan harus memberikan dampak positif dan keuntungan yang besar bagi sebuah instansi. Kegagalan dalam hal tersebut akan mempengaruhi kredibilitas seorang data engineer (Hoiles et al., 2017).

REFERENSI

- Alonso-Fernandez, C., Calvo-Morata, A., Freire, M., Martinez-Ortiz, I., & Fernández-Manjón, B. (2019). Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education, 141*, 103612.
- Arnott, D., Lizama, F., & Song, Y. (2017). Patterns of business intelligence systems use in organizations. *Decision Support Systems, 97*, 58–68.
- Cao, L. (2017). Data science: a comprehensive overview. *ACM Computing Surveys (CSUR), 50*(3), 1–42.
- Dinov, I. D. (2018). *Data Science and Predictive Analytics*. Springer.
- Dunn, P. F., & Davis, M. P. (2017). *Measurement and data analysis for engineering and science*. CRC press.
- Hao, J., & Ho, T. K. (2019). Machine learning made easy: A review of scikit-learn package in Python programming language. *Journal of Educational and Behavioral Statistics, 44*(3), 348–361.
- Hoiles, W., Aprem, A., & Krishnamurthy, V. (2017). Engagement and popularity dynamics of YouTube videos and sensitivity to meta-data. *IEEE Transactions on Knowledge and Data Engineering, 29*(7), 1426–1437.
- Kelleher, J. D., & Tierney, B. (2018). *Data science*. MIT Press.
- Richards, G., Yeoh, W., Chong, A. Y. L., & Popovič, A. (2019). Business intelligence effectiveness and corporate performance management: an empirical analysis. *Journal of Computer Information Systems, 59*(2), 188–196.
- Roy, S., Ray, R., Roy, A., Sinha, S., Mukherjee, G., Pyne, S., Mitra, S., Basu, S., & Hazra, S. (2017). IoT, big data science & analytics, cloud computing and mobile app based hybrid system for smart agriculture. *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON), 303–304*.
- Santika, I. G. N. (2017). Kepala Sekolah Dalam Konsep Kepemimpinan Pendidikan: Suatu Kajian Teoritis. *Widya Accarya, 7*(1). <http://103.39.12.42/index.php/widyaaccarya/article/view/898>
- Santika, I. G. N. (2020). Optimalisasi Peran Keluarga Dalam Menghadapi Persoalan Covid- 19 : Sebuah Kajian Literatur. *Jurnal Ilmiah Ilmu Sosial, 6*(2), 127–137. <https://doi.org/http://dx.doi.org/10.2388>

7/jiis.v6i2.28437

Santika, I. G. N. (2021). Grand Desain Kebijakan Strategis Pemerintah Dalam Bidang Pendidikan Untuk Menghadapi Revolusi Industri 4.0. *Jurnal Education and Development*, 9(2), 369–377.

Swarniti, N. W. (2019). The Translation Procedures of Bible Translation. *RETORIKA: Jurnal Ilmu Bahasa*, 5(2), 187–196.
<https://doi.org/10.22225/jr.5.2.1277.187-196>

Swarniti, N. W. (2021). A Corpus Based Approach to the Analysis of Structures

in Prepositional Phrase. *Yavana Bhasha: Journal of English Language Education*, 4(1), 18–22.

<https://doi.org/http://dx.doi.org/10.25078/yb.v4i1.2207>

Swarniti, N. W. (2021). Translation Methods Found in New Testament Bible of Mark 's Gospel. *RETORIKA: Jurnal Ilmu Bahasa*, 7(2), 172–179.

<https://doi.org/https://doi.org/10.22225/jr.7.2.3823.172-179>