

ALGORITMA SUPERVISED LEARNING DAN UNSUPERVISED LEARNING DALAM PENGOLAHAN DATA

Puput Santoso¹, Heri Abijono², Novita Lestari Anggreini³

¹Universitas AMIKOM Yogyakarta

²Sekolah Tinggi Teknologi Cahaya Surya

³Politeknik TEDC

agathapuput@gmail.com

ABSTRAK

Seiring dengan zaman yang semakin berkembang seperti saat ini, kini kita berada pada zaman yang mana teknologi menjadi satu hal yang paling penting dan tidak pernah terlepas dari kehidupan manusia. Dengan teknologi informasi dan komunikasi yang ada sekarang ini, semakin memudahkan kita untuk melakukan segala aktifitas. Dengan semakin berkembangnya teknologi informasi juga menghasilkan begitu banyak data yang dapat diolah, sehingga banyak informasi yang tidak terbuang sia-sia. Machine learning dapat digunakan sebagai sistem pengolahan data sehingga dapat mempermudah pengguna dalam mengolah informasi yang begitu banyak (big data), kemudian informasi ini dapat digunakan sebagai pedoman dalam berbagai bidang keilmuan yang membutuhkannya.

Kata kunci: Teknologi, Informasi, Komunikasi, Machine, Learning, Big Data

ABSTRACT

Along with the increasingly developing times like today, we are now in an age where technology is the most important thing and can never be separated from human life. With the information and communication technology that exists today, it is easier for us to carry out all activities. With the development of information technology also produces so much data that can be processed, so that a lot of information is not wasted. Machine learning can be used as a data processing system so that it can make it easier for users to process so much information (big data), then this information can be used as a guide in various scientific fields that need it.

Keywords: Technology, Information, Communication, Machine, Learning, Big Data

PENDAHULUAN

Kemajuan teknologi berkembang begitu pesat, dapat dikatakan bahwa dunia sedang mengalami era transformasi digital, yang mana sejak dari teknologi berukuran besar hingga teknologi berukuran mikro yang tidak pernah kita bayangkan di masa lalu, kini tercipta pada era ini. Dengan adanya teknologi, kegiatan sehari-hari manusia dapat dikerjakan dengan praktis, kemudian ditunjang peran Internet maka semakin memudahkan pekerjaan manusia, seluruh informasi mengenai apapun yang ingin diketahui dapat diakses melalui Internet baik informasi mengenai industri, politik, hiburan, dan lain-lain dapat ditemukan melalui Internet.

Dengan Internet, transaksi pun dapat dilakukan secara *online*, yang sekarang lebih dikenal

dengan nama *online shop*. *Online shop* bahkan dapat mengetahui kebutuhan yang sedang kita inginkan, sebab *online shop* dapat menyimpan data yang telah kita akses pada waktu sebelumnya, dan dengan adanya sistem *machine learning* yang dapat memberikan prediksi berdasarkan data yang telah ada pada waktu lalu (Louridas, 2016:110–115). *Machine learning* merupakan topik yang hangat di ranah teknologi beberapa tahun terakhir, karena dengan sistem ini diyakini akan mengubah serta mempermudah cara hidup dan cara kerja manusia. *Machine learning* dapat memprediksi dengan tingkat keakuratan yang tinggi, sehingga *machine learning* akan menjadi teknologi yang terpenting setelah Internet (Khairudin, 2018).

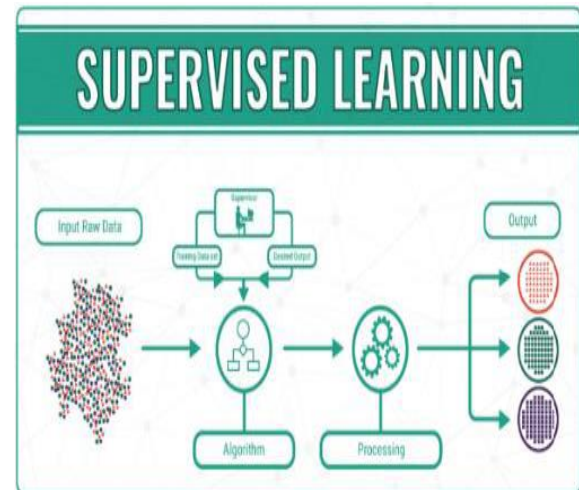
Machine learning muncul dengan salah satu manfaatnya adalah untuk dapat mengolah serta menganalisa *big data* (Chang, 2018:28–35). *Big data* merupakan jumlah data yang terkumpul pada tingkat global dengan jumlah yang tak terbayangkan. Terdapat potensi yang begitu besar pada *big data*, sehingga jika *big data* ini dapat diproses dan dianalisis dengan baik, maka dapat menghasilkan informasi yang sangat berguna untuk berbagai macam keperluan yang antara lain pada sektor bisnis. Setelah dilakukan analisis kemudian dapat diperoleh informasi seperti pengurangan biaya, pengembangan produk, pengambilan keputusan yang cerdas, dan masih banyak lagi informasi yang lain (Leary, 2013).

Penelitian ini bertujuan untuk memahami cara kerja algoritma *Supervised Learning* dan *Unsupervised Learning* dalam pengumpulan dan analisis data serta mengetahui tantangan akan pentingnya *machine learning* di masa yang akan datang.

METODE PENELITIAN

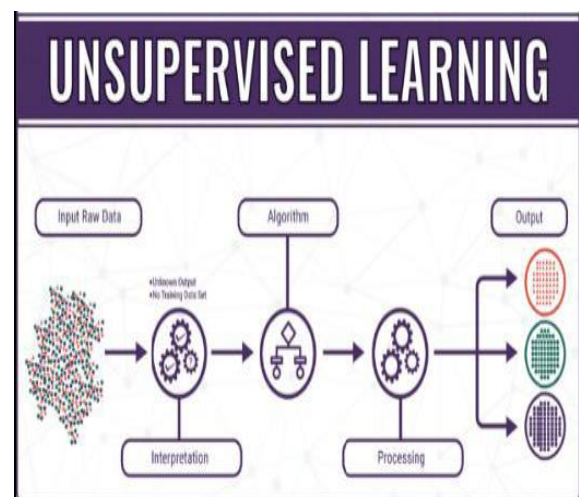
Beberapa percobaan dilakukan dengan menggunakan model *data set* yang berbeda. *Machine learning* memiliki tiga tipe, yaitu *Supervised Learning*, *Unsupervised Learning*, dan *Reinforcement Learning*. Masing-masing tipe tersebut di atas memiliki tipe *data set* yang berbeda pula. Pada penelitian ini melakukan percobaan dengan tipe *Supervised Learning* dan *Unsupervised Learning*.

Pada algoritma *Supervised Learning*, sistem diberikan *training data set* berupa informasi masukan dan keluaran yang diinginkan, sehingga sistem akan mempelajari berdasarkan data yang telah ada. Sistem akan mencari pola dari *data set*, kemudian pola itu akan dijadikan sebagai acuan untuk kumpulan data berikutnya. Gambar 1 menunjukkan diagram blok cara kerja *Supervised Learning*.



Gambar 1. Langkah Kerja *Supervised Learning*

Tahapan kerja pada algoritma *Unsupervised Learning* ditunjukkan pada Gambar 2.



Gambar 2. Langkah Kerja *Unsupervised Learning*

Algoritma *Unsupervised Learning* bersifat deskriptif, yang akan berguna untuk mengelompokkan atau mengkategorikan data. Algoritma ini tidak mendapatkan *training data set*, karena algoritma ini bukan bersifat prediktif, sehingga membutuhkan pembelajaran dari data yang telah ada.

Karena *machine learning* berfungsi untuk mengambil keputusan yang benar-benar akurat, *machine learning* menggunakan prinsip dari probabilitas dengan menggunakan berbagai macam kemungkinan yang akan terjadi, sehingga dengan adanya kemungkinan tersebut *machine learning* dapat menentukan keputusan yang akan diambil berdasarkan

kemungkinan yang dipelajari dari peristiwa yang telah ada sebelumnya (Advernesia, 2019).

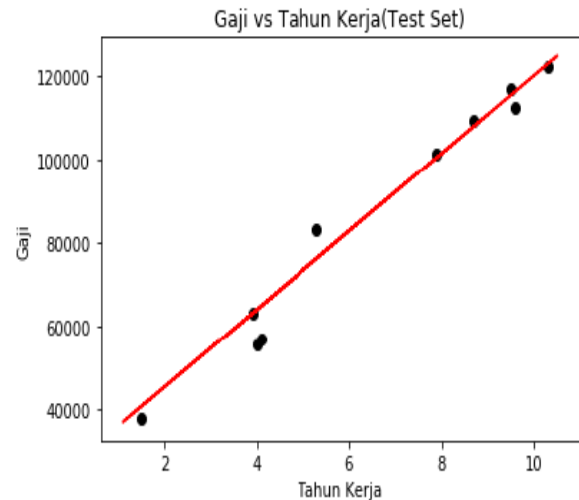
Probabilitas adalah kemungkinan-kemungkinan yang terjadi pada suatu peristiwa dengan memperkirakan kejadian apa yang akan terjadi. Maka dari itu probabilitas hanya memiliki rentang nilai antara 0 sampai 1, dengan nilai 0 menunjukkan bahwa peristiwa tidak terjadi, dan nilai 1 menunjukkan bahwa peristiwa tersebut terjadi, sehingga probabilitas dapat digunakan untuk meramalkan kejadian yang akan terjadi, yang kemudian kita bisa menentukan keputusan yang akan diambil (Murphy, 2012).

Statistika juga menjadi ilmu dasar dari *machine learning*. Dalam statistika dapat dilakukan pemahaman fenomena, penentuan keputusan, menguji hipotesis berdasarkan data. Statistika menyelesaikan masalah atau menentukan hasil dengan mencari pola yang memiliki nilai *error* terkecil, sehingga hasil yang ditemukan akurat. Statistika dapat digunakan dalam menentukan pola pada data, dan hal ini berbeda dengan *machine learning* yang dapat menentukan pola data dengan mengutamakan hasil yang optimal (Dangeti, 2017).

HASIL DAN PEMBAHASAN

Dari *data set* yang telah diberikan, *machine learning* menjalankan algoritma *Supervised Learning* dengan melakukan pendekatan dari satu data ke data yang lain. Pada bahasa pemrograman Python, pendekatan dilakukan dengan terlebih dahulu memecah data menjadi dua, yaitu 1) data yang akan dipelajari, dan 2) data yang akan dites terhadap data yang telah dipelajari. *Supervised learning* bersifat prediktif, sehingga dari banyak hasil pendekatan yang telah dilakukan akan didapatkan sebuah prediksi. Algoritma ini dapat menyelesaikan permasalahan data secara linear, multilinear, maupun polynomial.

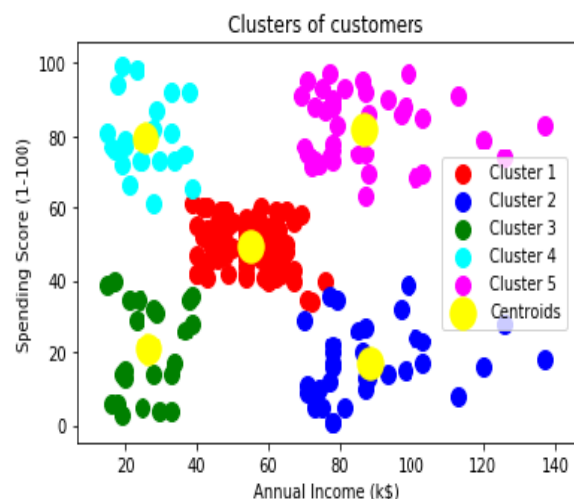
Sebagai contoh dengan menyelesaikan kasus pemberian gaji berdasarkan lamanya pengalaman kerja seseorang, kasus ini dapat diselesaikan dengan memberikan prediksi secara linear seperti yang ditunjukkan pada Gambar 3.



Gambar 3. Hasil Prediksi *Supervised Learning*

Garis merah pada Gambar 3 merupakan hasil prediksi, dan titik-titik berwarna hitam merupakan beberapa pengtesan yang diambil dari pemecahan data. Dapat dilihat bahwa sebagian besar dari nilai yang dites berada pada garis prediksi, sehingga hasil prediksi dapat dikatakan baik.

Untuk algoritma *Unsupervised Learning*, penulis melakukan percobaan dengan data acak mengenai berbagai pelanggan dengan pendapatan gaji yang beragam, dan kemudian algoritma *Unsupervised Learning* melakukan pengelompokkan menjadi lima kategori. Gambar 4 menunjukkan hasil kerja algoritma ini dengan data yang diberikan.



Gambar 4. Hasil Pengolahan pada *Unsupervised Learning*

Gambar 4 menunjukkan bahwa data awal telah dikelompokkan menjadi lima *cluster* berbeda. *Unsupervised Learning* ini memiliki sifat

deskriptif, maka algoritma ini hanya menunjukkan hasil proses seperti pada Gambar 4 di atas.

KESIMPULAN DAN SARAN

Kesimpulan

Machine learning dipakai untuk mencari pola pada data. Jika pola dari data sudah diketahui, maka *machine learning* dapat mengambil kesimpulan berupa pola yang telah terbentuk, sehingga dapat menentukan keputusan setelah menentukan kesimpulan.

Machine learning memiliki sifat prediktif dan deskriptif yang membuktikan bahwa *machine learning* akan menjadi suatu problematika yang amat penting di masa mendatang, karena dengan *machine learning* pengguna dapat menentukan keputusan dengan tepat.

Saran

Machine learning dapat berfungsi untuk menyelesaikan permasalahan *big data*, sehingga *machine learning* merupakan hal

yang krusial demi memasuki revolusi industri 4.0. Maka dari itu pendalaman tentang *machine learning* perlu terus diterapkan.

UCAPAN TERIMAKASIH

Pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada pihak-pihak yang telah membantu di dalam penyusunan dan penerbitan makalah ini, yaitu kepada yang terhormat: (1) Henri Tetiawadi, M.M., M.Kom., yang telah mengizinkan penulis menempuh studi Strata-2 pada program studi Magister Teknik Informatika pada Universitas AMIKOM Yogyakarta, (2) Para pengurus *G-Tech Journal* UNIRA Malang yang telah mempublikasikan makalah ilmiah ini, dan (3) Heri Abijono, M.Kom. dan Novita Lestari Anggreini, S.ST. M.Kom. yang telah bersedia bekerja sama di dalam melakukan praktikum dan *literature review* sehingga makalah ilmiah ini dapat disusun.

REFERENSI

- Advernesia. (2019). Apa Itu Machine Learning dan Cara Kerjanya. https://www.advernesia.com/blog/data-science/machine-learning-adalah/#llc_comments. Diakses pada tanggal 20 Desember 2020.
- Chang, Z., Lei, L., Zhou, Z., Mao, S., & Ristaniemi, T. (2018). Learn to Cache : Machine Learning for Network Edge Caching in the Big Data Era. *IEEE Wirel. Commun.* Vol. 25. hlm. 28–35.
- D. E. O. Leary. (2013). *Artificial Intelligence and Big Data*. California.
- K.P. Murphy. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press.
- Khairudin, I. (2018). Machine Learning Akan Jadi Teknologi Terpenting Setelah Internet di 2018. Selular. ID. <https://selular.id/2018/01/machine-learning-akan-menjadi-teknologi-terpenting-setelah-internet-di-2018/>. diakses pada tanggal 12 Januari 2021.
- Louridas, P. & Ebert, C. (2016). Machine Learning. *IEEE Softw.* Vol. 33. hlm. 110–115.
- P. Dangeti. (2017). *Statistics for Machine Learning*. Packt, Mumbai.