

Bagian 2 :

1. *Outlier* adalah suatu nilai yang dianggap tidak normal. *Outlier* atau yang biasa disebut juga dengan nilai pencilan bisa dikatakan nilai tersebut mempunyai perbedaan yang sangat jauh dengan pusat data. Sedangkan menurut (Rosa et al., 2015), *outlier* merupakan data yang dianggap tidak mirip atau tidak konsisten dengan data lainnya. *Outlier* atau nilai pencilan ini dapat menyebabkan distorsi terhadap nilai yang asli. Implikasi nilai *outlier* atau pencilan pada analisis data akan mengakibatkan ketidaksesuaian data analisa dan ketidakakuratan model *machine learning* yang sedang dikerjakan. Cara manajemen data untuk menangani kasus *outlier* ini bisa dengan mengganti nilai *outliers* menggunakan nilai *lower bound* dan *upper bound*, atau bisa juga dengan menggantinya dengan nilai mean, modus atau median. (Sumber: sanbercode)
2. Korelasi atau dapat diartikan juga sebagai hubungan, dalam kamus besar bahasa indonesia artinya adalah hubungan timbal balik atau sebab akibat. Korelasi bisa juga untuk menentukan derajat asosiasi menurut (Asuero et al., 2006) Pada dunia statistika, korelasi adalah salah satu tehnik yang digunakan untuk mencari hubungan 2 variabel yang bersifat kuantitatif. Implikasi korelasi dalam teori statistika adalah untuk mengukur atau mengetahui seberapa kuat hubungan antar variabel. Manfaat dari mengetahui seberapa kuat hubungan antar variabel tersebut bisa untuk mendeskripsikan gambaran yang lebih bermanfaat dari data-data yang ada. Salah satu contohnya yaitu suatu perusahaan menggunakan korelasi dalam statistika untuk menentukan hubungan antara variabel pendapatan perusahaan dan variabel gaji karyawan. Dari analisis korelasi antara kedua variabel tersebut, maka perusahaan tersebut dapat membuat suatu keputusan.
3. *Machine learning* atau pembelajaran mesin merupakan salah satu cabang dari ilmu kecerdasan buatan yang dikembangkan untuk bisa belajar sendiri tanpa arahan dari penggunanya. Ilmu statistika, matematika, dan *data mining* diperlukan dalam *machine learning* untuk membuat sistem dapat belajar dan menganalisa sendiri tanpa perlu diprogram ulang atau diperintah. Meskipun dari disiplin ilmu statistika dan *machine learning* kelihatan sangat berbeda, namun dua disiplin ilmu ini ada kaitannya. *Machine learning* dan statistika adalah dua hal yang saling melengkapi (Cunningham, n.d.). Jadi untuk membuat *machine learning* diperlukan ilmu dari statistika untuk membuat peluang, pemodelan untuk membuat prediksi data baru. Sebaliknya, statistika juga dapat menggunakan machine learning untuk mempermudah pekerjaan yang dapat digunakan untuk mengambil kesimpulan dari data yang digunakan.
4. Pada dasarnya *artificial intelligence*, *machine learning*, dan *deep learning* adalah saling berkaitan (Trijaya, n.d.). *Artificial intelligence* atau kecerdasan buatan merupakan gabungan dari beberapa cabang, salah satunya yaitu *machine learning* atau pembelajaran mesin. Sedangkan *machine learning* sendiri mempunyai pecabangan diantaranya adalah neural network. Neural network atau jaringan syaraf tiruan tersebut merupakan teknik dalam *machine learning* yang menirukan syaraf pada manusia. Syaraf sendiri adalah bagian fundamental dari otak. Dalam neural network tersebut ada percabangan teknik yang dinamakan deep learning. Jadi Deep learning adalah teknik percabangan dari Neural network, sedangkan neural network merupakan percabangan dari *machine learning*, dan *machine learning* sendiri adalah teknik percabangan dari *artificial intelligence* atau kecerdasan buatan.

5. Intrepetasi data adalah menjelaskan kepada orang lain mengenai hasil dari pengolahan data yang kita dapatkan melalui sudut pandang penafsir. Interpretasi data bisa juga diartikan sebagai proses meninjau data melalui beberapa proses yang telah ditentukan sebelumnya yang akan membantu membuat kesimpulan. Namun menurut ahli Intrepetasi data merupakan deskripsi dan ekspresi dalam suatu penelitian kualitatif. Intepretasi data juga dapat diartikan sebagi usaha untuk menggali pengetahuan tentang data atau peristiwa melalui pemikiran yang lebih mendalam. Tantangan dari intreptasi data yaitu seseorang yang akan menafsirkan atau mengintrepetasikan data seharusnya mempunyai dasar pengetahuan yang kuat dan pengalaman yang matang agar dapat memahami dan mengambil makna dari hasil penafsiran atau intrepetasi tersebut. Intreptasi data dengan *storry telling* dan *decision making* sangat erat hubungannya. Kaitannya dengan *story telling* yaitu seseorang yang akan mengintrepetasi data harus bisa menjelaskan data yang dihasiljan kepada orang lain dengan baik agar orang lain dapat mudah memahami dan tidak salah paham dari apa yang dijelaskan. Setelah menjelaskan kepada orang lain atau *story telling*, hal yang perlu dilakukan yaitu mengambil keputusan atau *decision making*. Pengambilan keputusan diperlukan dalam intrepetasi data untuk mengambil sebuah kesimpulan. Setelah menjelaskan mengenai data yang dihasilkan, langkah selanjutnya yaitu mengambil kesimpulan. Dari kesimpulan tersebut bisa digunakan untuk mengambil keputusan atau *decision making*.

Asuero, A. G., Sayago, A., & González, A. G. (2006). The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry*, 36(1), 41–59.
<https://doi.org/10.1080/10408340500526766>

Cunningham, S. J. (n.d.). *Machine Learning and Statistics: A matter of perspective*.

Rosa, P. H. P., Kwa, F. E., & Rosa, P. H. P. (2015). *Deteksi Outlier Menggunakan Algoritma Block- based Nested Loop (Studi Kasus : Data Akademik Mahasiswa ... Deteksi Outlier Menggunakan Algoritma Block-based Nested Loop (Studi Kasus : Data Akademik Mahasiswa Prodi PS Universitas XYZ). March*, 3–6.

Trijaya, T. A. M. (n.d.). *Mengenal Artificial Intelligence, Machine Learning, & Deep Learning*. <https://amt-it.com/mengenal-perbedaan-artificial-intelligence-machine-learning-deep-learning/>

<https://www.formpl.us/blog/data-interpretation>