

Nama anggota kelompok:

1. Safira Amalia Puspitasari R (2101191008)
2. Agita Purwandani (2101191006)
3. Tyas Oksi Praktika (2101191014)

TUGAS BESAR KELOMPOK 11

ANALISIS NEURAL NETWORK DENGAN DATA SET SEEDS

Neural Network merupakan kategori ilmu Soft Computing. Neural Network sebenarnya mengadopsi dari kemampuan otak manusia yang mampu memberikan stimulasi/rangsangan, melakukan proses, dan memberikan output. Output diperoleh dari variasi stimulasi dan proses yang terjadi di dalam otak manusia. Kemampuan manusia dalam memproses informasi merupakan hasil kompleksitas proses di dalam otak.

1. Struktur Neural Network

Konsep dasar pembangunan neural network buatan (*Artificial Neural Network*) terbentuk. Ide mendasar dari Artificial Neural Network (ANN) adalah mengadopsi mekanisme berpikir sebuah sistem atau aplikasi yang menyerupai otak manusia, baik untuk pemrosesan berbagai sinyal elemen yang diterima, toleransi terhadap kesalahan/*error*, dan juga *parallel processing*. *Neural network* dibangun dari banyak node/unit yang dihubungkan oleh link secara langsung. Link dari unit yang satu ke unit yang lainnya digunakan untuk melakukan propagasi aktivasi dari unit pertama ke unit selanjutnya. Setiap link memiliki bobot numerik. Bobot ini menentukan kekuatan serta penanda dari sebuah konektivitas.

Proses pada ANN dimulai dari input yang diterima oleh neuron beserta dengan nilai bobot dari tiap-tiap input yang ada. Setelah masuk ke dalam neuron, nilai input yang ada akan dijumlahkan oleh suatu fungsi perambatan (*summing function*), yang bisa dilihat seperti pada di gambar dengan lambang sigma (Σ). Hasil penjumlahan akan diproses oleh fungsi aktivasi setiap neuron, disini akan dibandingkan hasil penjumlahan dengan *threshold* (nilai ambang) tertentu. Jika nilai melebihi *threshold*, maka aktivasi neuron akan dibatalkan, sebaliknya, jika masih dibawah nilai *threshold*, neuron akan diaktifkan. Setelah aktif, neuron

akan mengirimkan nilai output melalui bobot-bobot outputnya ke semua neuron yang berhubungan dengannya. Proses ini akan terus berulang pada input-input selanjutnya.

1.1 Data Set

Data yang digunakan dalam pengerjaan Tugas Besar ini berupa data .csv yang didapatkan dari website kaggle yaitu <https://www.kaggle.com/rwzhang/seeds-dataset>. Data ini merupakan data ekstrasi ciri dari *seeds* dan memiliki 210 data dan didalamnya memiliki tiga kelas yaitu Kama(1), Rosa(2) dan Canadian(3) dan tujuh fitur yaitu area, perimeter, compactness, length of kernel, width of kernel, asymmetry coefficient, length of kernel groove. Masing-masing kelas memiliki 70 data.

2. ANALISIS HASIL PERANCANGAN SISTEM

2.1. Analisis Skenario Pertama

Berdasarkan pengambilan data yang dibagi menjadi dua bagian yaitu data uji dan data latih dengan data uji 20% dan data latih 80% selanjutnya diacak dengan nilai 4 yang ditunjukkan pada gambar 1. Pengukuran dan pengujian scenario pertama ini menggunakan data uji latih sebanyak 168 data, parameter dua hidden layer, pada layer pertama terdiri atas 10 neuron, layer kedua terdiri dari 5 neuron maksimum iterasi 700 kali, selebihnya parameter yang digunakan adalah auto dari mlpclassifer.

	precision	recall	f1-score	support
1	0.95	0.95	0.95	58
2	0.98	0.98	0.98	55
3	0.96	0.96	0.96	55
accuracy			0.96	168
macro avg	0.96	0.96	0.96	168
weighted avg	0.96	0.96	0.96	168

Gambar 1. Akurasi pada algoritma NN

Dapat dilihat pada gambar 1. akurasi pada algoritma NN dapat berubah ubah dikarenakan bobotnya berubah-ubah, maka dari itu kami melakukan training sebanyak 5 kali dan merata-rata hasil tersebut

Training 1 = 93%

Training 2 = 97%

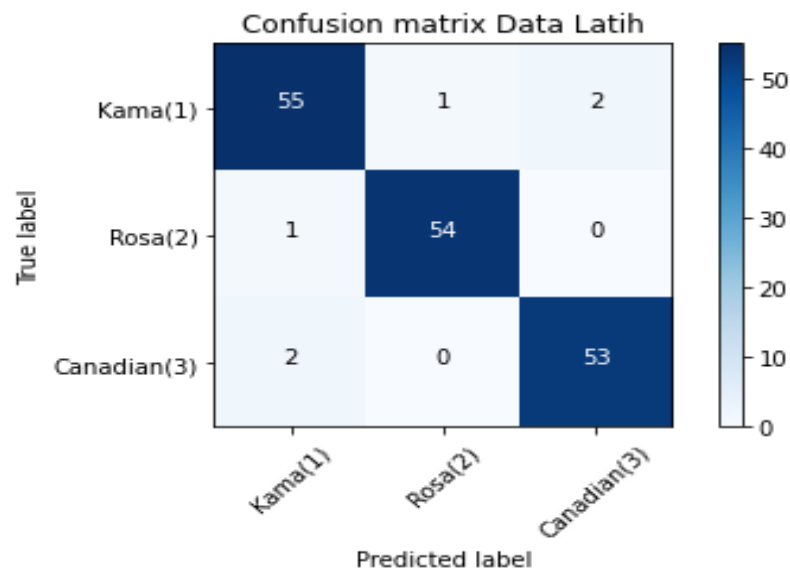
Training 3 = 94%

Training 4 = 98%

Training 5 = 96%

Rata-rata = 95.6%

Matriks yang sudah ada lalu di proses dengan melakukan plot *confusion matrix* pada setiap kelas/kategori, tujuannya untuk mengetahui kelas mana yang masuk kedalam kelas lain, dapat dilihat pada gambar 2.



Gambar 2. Confusion Matrix Data Latih

Dapat dilihat pada figure diatas saat training pertama bahwa kelas Kama(1) sebanyak 4 data diprediksi bahwa itu adalah kelas Canadian(3). Saat training kedua bahwa kelas Canaadian(3) sebanyak 2 data diprediksi bahwa itu adalah kelas Kama(1). Saat training ketiga bahwa kelas Canaadian(3) sebanyak 2 data diprediksi bahwa itu adalah kelas Kama(1)

dan kelas Kama(1) sebanyak 4 data diprediksi bahwa itu adalah kelas Canadian(3). Saat training keempat bahwa kelas Canadian(3) sebanyak 2 data diprediksi bahwa itu kelas Kama(1) dan sebanyak 2 data Kama(1) diprediksi sebagai Canadian(3). Saat training kelima hasil sama dengan training keempat namun ada satu data Rosa(2) yang diprediksi sebagai Kama(1) dan sebaliknya

2.2. Analisis Skenario kedua

Scenario kedua digunakan untuk mendapatkan akurasi pengujian yang diambil dari data uji dengan jumlah 42 data, fitur yang digunakan sebanyak 7 ciri, tiap fitur memiliki nilai rata-rata 0 dan variasi 1. Sistem klasifikasi Neural Network ini menggunakan parameter dua hidden layer dengan hidden layer pertama 10 neuron dan hidden layer kedua 5 neuron dan maksimum iterasi 700 kali, selebihnya paramater yang digunakan adalah auto dari mlpclassifer, dapat dilihat pada gambar 3.

	precision	recall	f1-score	support
1	0.92	1.00	0.96	12
2	1.00	1.00	1.00	15
3	1.00	0.93	0.97	15
accuracy			0.98	42
macro avg	0.97	0.98	0.98	42
weighted avg	0.98	0.98	0.98	42

Gambar 3. Akurasi pengujian menggunakan data uji.

Pada hasil pengujian data uji dapat dilihat hasil dari akurasi pengujian, akurasi pengujian diperlakukan sama dengan pegujian data latih yaitu dengan melakukan rata-rata akurasi dari 5 kali pengujian.

Pengujian 1 = 97%

Pengujian 2 = 97%

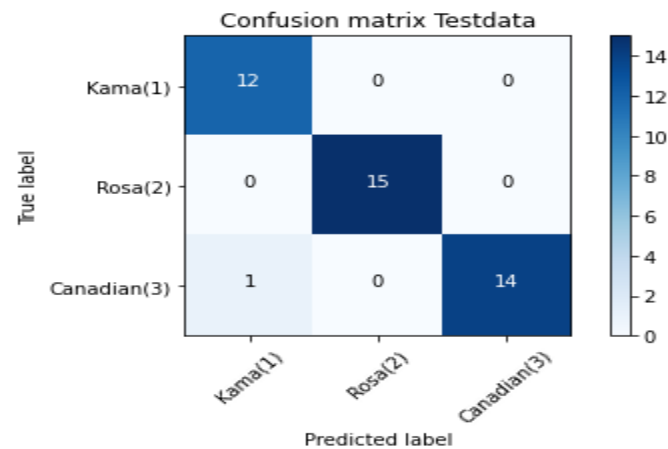
Pengujian 3 = 97%

Pengujian 4 = 97%

Pengujian 5 = 97%

Rata-rata = 97%

Selanjutnya data tersebut di citrakan menggunakan plot *confusion matrix* pada setiap kelas/kategori. Dapat dilihat pada gambar 4



Gambar 4. Confution Matrik Data Uji

- Pada hasil pengujian pertama dapat dilihat kelas Canadian(3) ada satu data yang masuk prediksi pada kelas Kama(1)
- Pada hasil pengujian kedua hasil sama dengan pengujian pertama
- Pada hasil pengujian ketiga hasil sama dengan pengujian pertama dan kedua
- Pada hasil pengujian keempat hasil sama dengan pengujian sebelumnya
- Pada hasil pengujian kelima hasil sama dengan pengujian sebelumnya

KESIMPULAN

Akurasi pada algoritma NN dapat berubah ubah dikarenakan bobotnya berubah-ubah dari 5 kali percobaan, dimana nilai training tertinggi pada uji coba ke 4 yaitu 98%, rata rata akurasi pelatihan pada lagoritma NN adalah 95.6%. Hasil pengujian data uji dapat dilihat hasil dari akurasi pengujian dengan 5 kali percobaan mendapatkan hasil yang sama yaitu 97%. Dilihat dari data diatas bahwa kelas Kama(1) diprediksi menjadi kelas Candian(3) dan kebalikannya, kemungkinan dikarenakan nilai salah satu ciri fitur dari kedua kelas tersebut hampir mirip.

