

## BAB IV

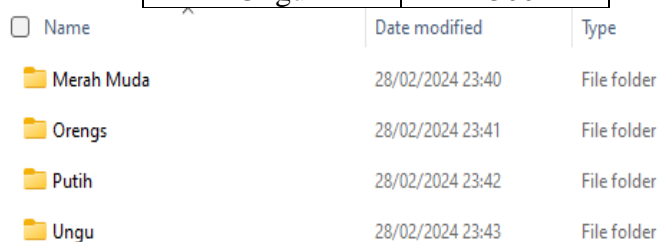
### IMPLEMENTASI SISTEM

#### 4.1 Dataset *Bougainvillea*

Penelitian ini menggunakan dataset citra bunga kertas (*bougainvillea*) dengan jumlah 5200 gambar, kemudian di kelompokkan menjadi 4 kelas dengan membuat 4 folder yaitu merah muda, *orengs*, putih dan ungu. Berikut Tabel 4.1 Jumlah Data dari Setiap Kelas dan Gambar 4.1 Dataset Citra Bunga Kertas (*Bougainvillea*).

Tabel 4.1 Jumlah Data dari Setiap Kelas

Kelas	Jumlah
Merah muda	1300
Orengs	1300
Putih	1300
Ungu	1300



Name	Date modified	Type
Merah Muda	28/02/2024 23:40	File folder
Orengs	28/02/2024 23:41	File folder
Putih	28/02/2024 23:42	File folder
Ungu	28/02/2024 23:43	File folder

Gambar 4.1 Dataset Citra Bunga Kertas (*Bougainvillea*)

#### 4.2 Skenario Pembagian Data

Pembagian sampel citra. Selanjutnya data di bagi menjadi dua kelompok data, yang akan digunakan untuk data latih dan data uji. Penelitian ini menerapkan lima skenario pembagian data, seperti yang di sajikan pada Tabel 4.2 Skenario pembagian data latih dan uji dan dari kelima pembagian data tersebut yang lebih unggul terdapat pada pembagian data 90% : 10% yang dapat di lihat pada Tabel 4.3 Pembagian sampel citra.

Tabel 4.2 Skenario pembagian data latih dan uji

Skenario	Jumlah Data Latih	Jumlah Data Uji
1	50%=2600	50%=2600
2	60%=3120	40%=2080
3	70%=3640	30%=1560
4	80%=4160	20%=1040
5	90%=4680	10%=520

Dari skenario pembagian data pada Tabel 4.2 Skenario pembagian data latih dan uji diperoleh akurasi dari masing-masing perbandingan dapat dilihat pada Tabel 4.3 Perolehan Akurasi perbandingan. Pada perbandingan akurasi tersebut di peroleh perbandingan terungul yang di gunakan untuk klasifikasi citra bunga kertas (*bougainvillea*) adalah 90% : 10% .

Tabel 4.3 Perolehan Akurasi perbandingan

No	Perbandingan	Akurasi <i>Neural Network</i>	Akurasi <i>Random Forest</i>
1	50% : 50%	99,6%	91,5%
2	60% : 40%	99,7%	94%
3	70% : 30%	99,7%	94,4%
4	80% : 20%	99,8%	94%
5	90% : 10%	99,8%	94,9%

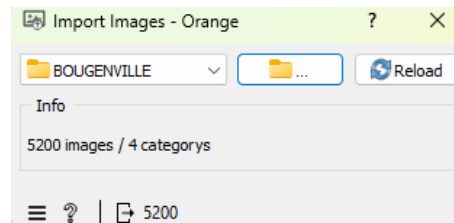
### 4.3 Widget Dalam Orange Data Mining

#### 4.3.1 Import Image

Kumpulan data yang telah di kelompokkan ke dalam kelas masing-masing direktori. Dimana pada tahap ini dilakukan import seluruh dataset *bougainvillea* kedalam aplikasi *orange* menggunakan widget *Import Image*. Dataset yang di *import* sebanyak 5200 citra dan 4 kategori dapat di lihat pada Gambar 4.2 widget *Import Image* dan Gambar 4.3 Tampilan widget *Import Image*.



Gambar 4.2 *widget Import Image*



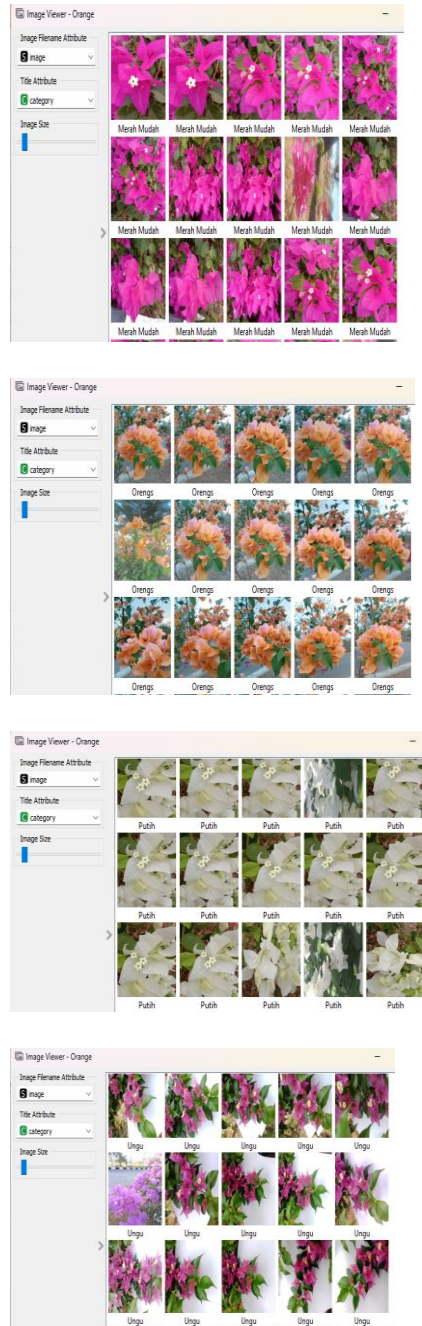
Gambar 4.3 Tampilan *widget Import Image*

#### 4.3.2 *Image Viewer*

*Image Viewer* adalah komponen dalam Orange yang digunakan untuk menampilkan gambar secara interaktif atau memungkinkan pengguna untuk memuat gambar dari berbagai format dan sumber, termasuk file lokal, URL, atau bahkan dari *dataset Orange*. Memastikan dataset telah berhasil di impor maka dibutuhkan *widget Image Viewer*. *widget Image Viewer* dan tampilan gambar pada *dataset bougainvillea* dapat di liat pada Gambar 4.4 *widget Image Viewer* dan Gambar 4.5 Tampilan *widget Image Viewer* .



Gambar 4.4 *widget Image Viewer*



Gambar 4.5 Tampilan *widget Image Viewer*

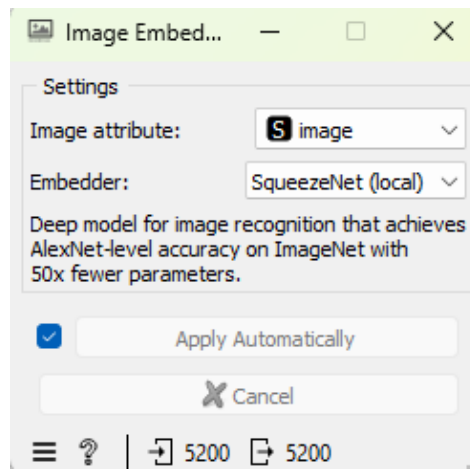
### 4.3.3 *Image Embedding*

Tahap ini dilakukan proses *image embedding* pada *Orange Data Mining* menggunakan *squeezeNet(local)* ini sebagai ekstraksi fitur pada citra. Pada tahap ini, *squeezeNet(local)* adalah sebuah arsitektur jaringan saraf tiruan (*Neural Network*) yang dikembangkan untuk tugas klasifikasi gambar.

*SqueezeNet (local)* untuk *Random Forest*, sebagai ekstraktor fitur untuk mengekstraksi ciri-ciri gambar, dan kemudian menggunakan ciri-ciri yang diekstraksi tersebut sebagai input untuk model *Random Forest*.. Hasil *embedding* ini nantinya akan digunakan sebagai data untuk klasifikasi menggunakan *Neural Network* dan *Random Forest*. Proses ini dapat dilihat pada Gambar 4.6 *widget image embedding* dan Gambar 4.7 Tampilan *widget image embedding* .



Gambar 4.6 *widget image embedding*

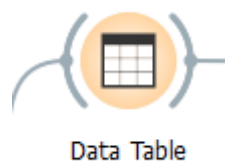


Gambar 4.7 Tampilan *widget image embedding*

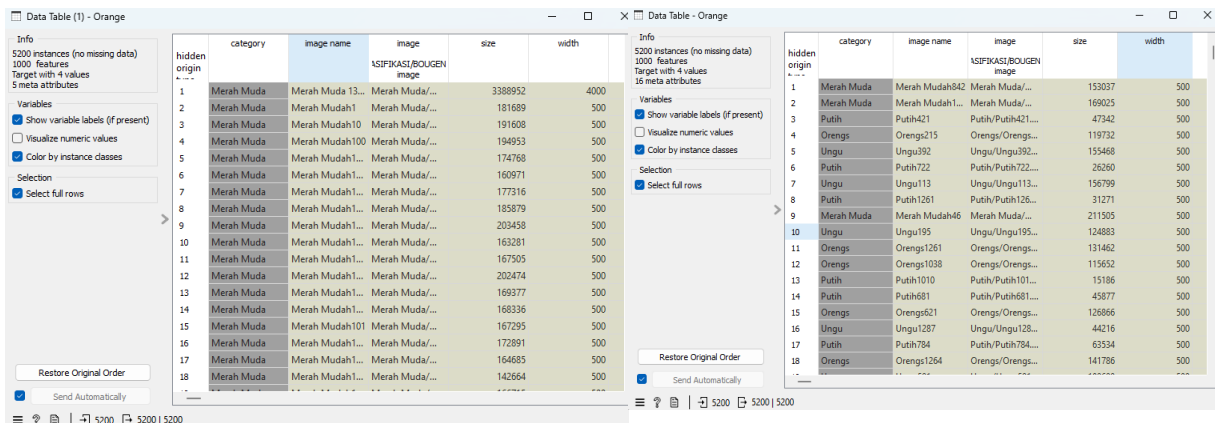
#### 4.3.4 Data Table

*Data table* adalah representasi visual dari data dalam bentuk tabel. Tabel ini memungkinkan pengguna untuk melihat, menyaring, dan menganalisis data secara lebih terperinci. Setiap baris dalam tabel mewakili satu instansi atau sampel data, sementara setiap kolom mewakili atribut atau fitur

dari data tersebut. *Data table* dalam penelitian ini terdapat dua *data table* yaitu *Data Table* sebelum di olah (*info 5200 instances (no missing data), 1000 features, target with 4 valuens dan 5 meta attributes*) dan *Data Table1* setelah di olah (*info 5200 instances (no missing data), 1000 features, target with 4 valuens dan 16 meta attributes*). Proses ini dapat di lihat pada Gambar 4.8 *widjet Data table* dan Gambar 4.9 Tampilan *widjet Data table* .



Gambar 4.8 *widjet Data table*



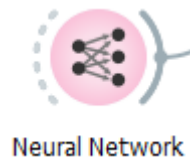
*Data Table* sebelum di olah

*Data Table* sesudah di olah

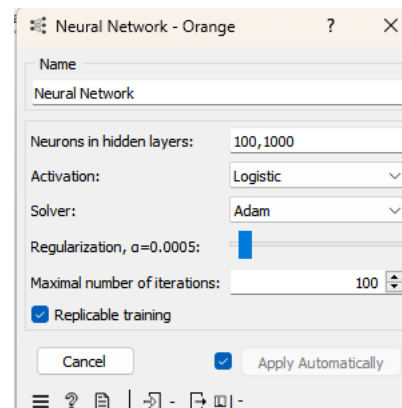
Gambar 4.9 Tampilan *widjet Data table*

### 4.3.5 Neural Network

Pada tahap ini ditambahkan *widjet Neural Network* dengan beberapa parameter yang di setting. Adapun parameter *Neural Network* yang digunakan dapat dilihat pada Gambar 4.10 *widjet Neural Network* dan Gambar 4.11 Tampilan *widjet Neural Network* .



Gambar 4.10 *widget Neural Network*



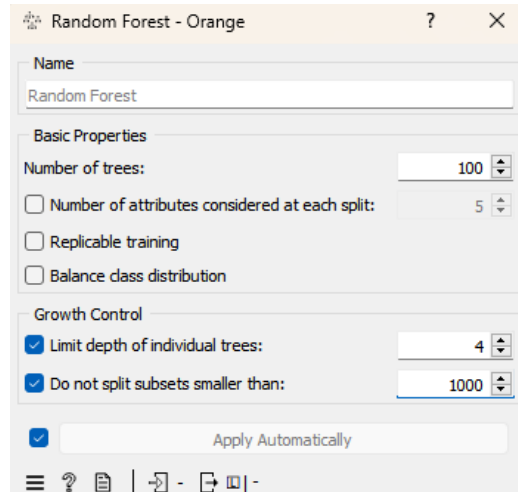
Gambar 4.11 Tampilan *widget Neural Network*

#### 4.3.6 Random Forest

*Random Forest* adalah sebuah algoritma *Machine Learning* yang digunakan untuk tugas-tugas seperti klasifikasi dan regresi. Ini adalah jenis algoritma yang termasuk dalam kategori ensemble learning, yang berarti itu menggabungkan prediksi dari beberapa model pembelajaran mesin yang lebih sederhana untuk menghasilkan prediksi yang lebih kuat. Proses ini dapat dilihat pada Gambar 4.12 *widget Random Forest* dan Gambar 4.13 Tampilan *widget Random Forest*.



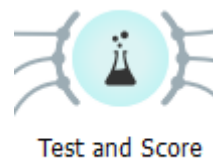
Gambar 4.12 *widget Random Forest*



Gambar 4.13 Tampilan *widget Random Forest*

#### 4.3.7 *Test and Score*

*Test and score* digunakan untuk mengevaluasi kinerja model prediksi yang telah dibuat. Ini memungkinkan pengguna untuk menguji model prediksi mereka dengan menggunakan data yang belum dilihat sebelumnya dan menilai seberapa baik model tersebut dapat memprediksi nilai target atau kelas. Dalam penelitian ini menggunakan dua bagian dalam *test and score* dapat di lihat pada Gambar 4.14 *widget test and score* dan Gambar 4.15 Tampilan *widget test and score*.



Gambar 4.14 *widget test and score*

Model	AUC	CA	F1	Prec	Recall	MCC
<i>Neural Network</i>	1	0,998	0,998	0,998	0,998	0,998
<i>Random Forest</i>	0,995	0,953	0,953	0,953	0,953	0,938

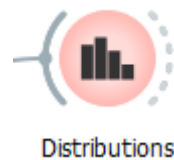


Gambar 4.15 Tampilan *widget test and score*

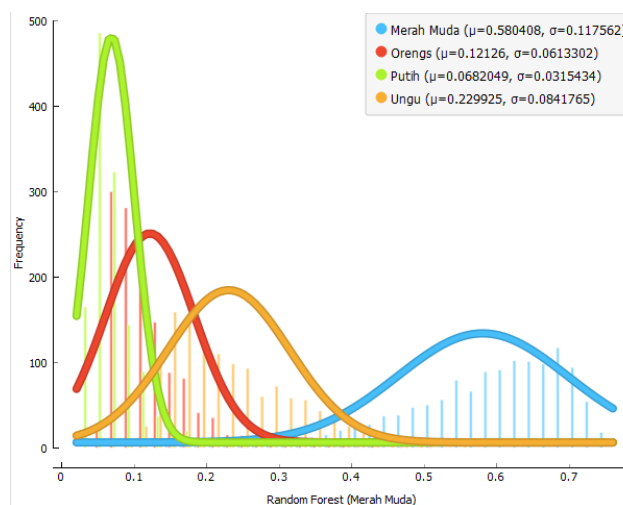
Dalam penelitian ini menggunakan *20-fold cross validation* dan *random sampling* untuk melakukan perbandingan dan pengujian data .

#### 4.3.8 *Distributions*

Dalam beberapa algoritma prediktif, seperti dalam analisis regresi linear, asumsi dasarnya adalah bahwa residu (selisih antara nilai prediksi dan nilai sebenarnya) memiliki distribusi normal. Dengan menggunakan kurva normal, dapat memeriksa seberapa baik asumsi ini dipenuhi oleh dataset. Dalam penelitian ini menggunakan *variable size* dan *fitted distribution normal* dapat di lihat pada Gambar 4.16 *widget distributions* dan Gambar 4.17 Tampilan *widget distributions*.



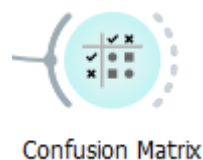
Gambar 4.16 *widget distributions*



Gambar 4.17 Tampilan *widget distributions*

#### 4.3.9 Confusion matrix

Terdapat suatu konsep dalam *orange data mining* dimana penghitungan akurasi dapat menggunakan salah satu metode yaitu *confusion matrix*. Langkah ini dilakukan perhitungan kinerja algoritma *Neural Network* dan *Random Forest* dengan menggunakan *widget confusion matrix* untuk memperlihatkan presentasi hasil dari setiap kelas sebagaimana dapat dilihat pada Gambar 4.18 *widget confusion matrix* dan Gambar 4.19 Tampilan *widget confusion matrix*, untuk memperlihatkan hasil evaluasi dari *test and score* yang mencakup jumlah instan antara kelas prediksi dan kelas aktual.



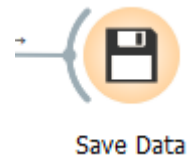
Gambar 4.18 *widget confusion matrix*

		Predicted				$\Sigma$
		Merah Muda	Orengs	Putih	Ungu	
Actual	Merah Muda	1228	31	0	41	1300
	Orengs	12	1256	30	2	1300
	Putih	0	26	1274	0	1300
	Ungu	58	34	10	1198	1300
$\Sigma$	1298	1347	1314	1241	5200	

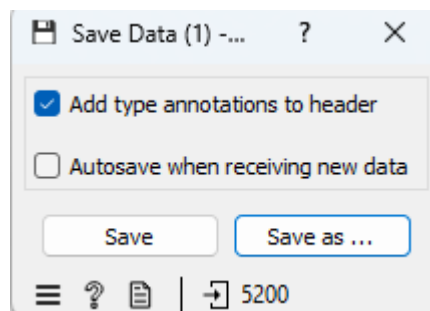
Gambar 4.19 Tampilan *widget confusion matrix*

#### 4.3.10 *Save Data*

*Save data* merupakan proses penyimpanan dataset atau hasil dari proses analisis ke dalam format yang dapat di akses dan digunakan kembali. Penelitian ini menggunakan format excel untuk menyimpan file analisis yang tela di lakukan dalam *orange data mining* dapat di lihat pada Gambar 4.20 *widget Save data* dan Gambar 4.21 Tampilan *widget Save data*.



Gambar 4.20 *widget Save data*



Gambar 4.21 Tampilan *widget Save data*