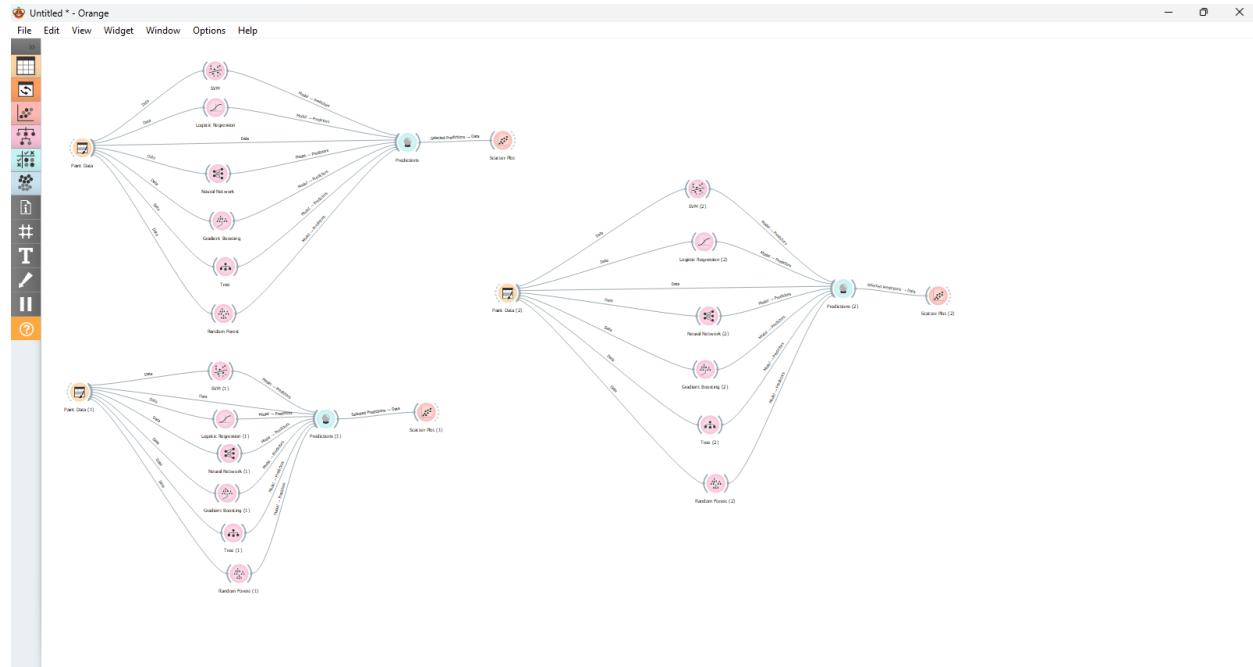
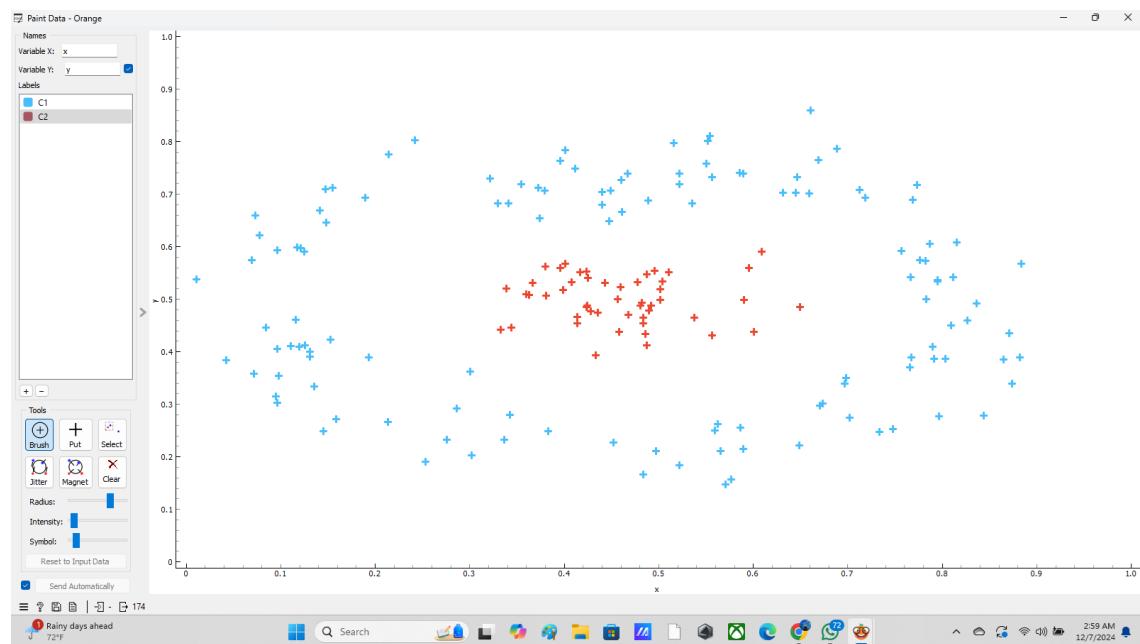


Nama : Ketut Satria Wibisana  
NIM : 1103213148

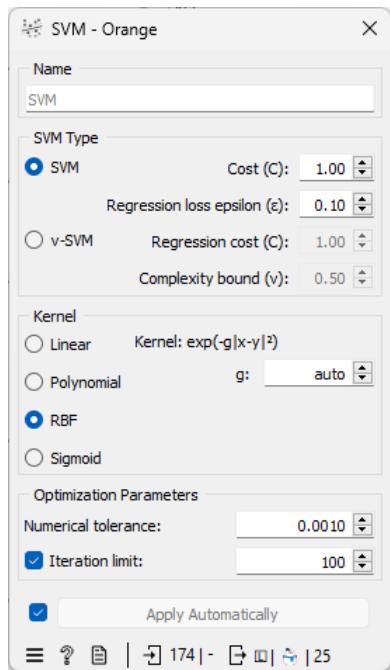
## WEEK 10 ORANGE DATA MINING



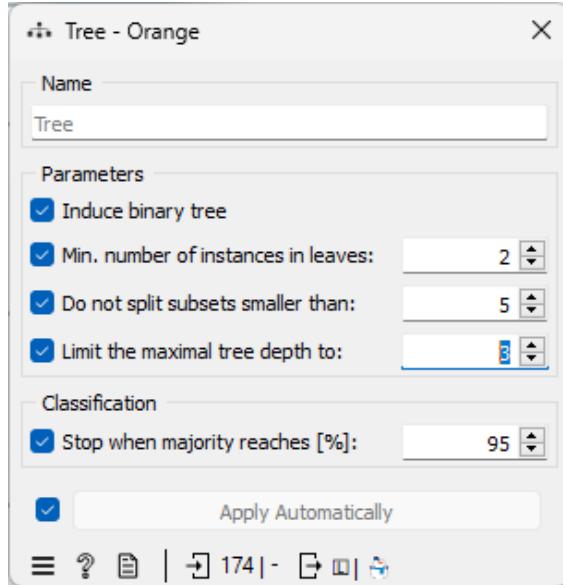
Paint Pola data 1:



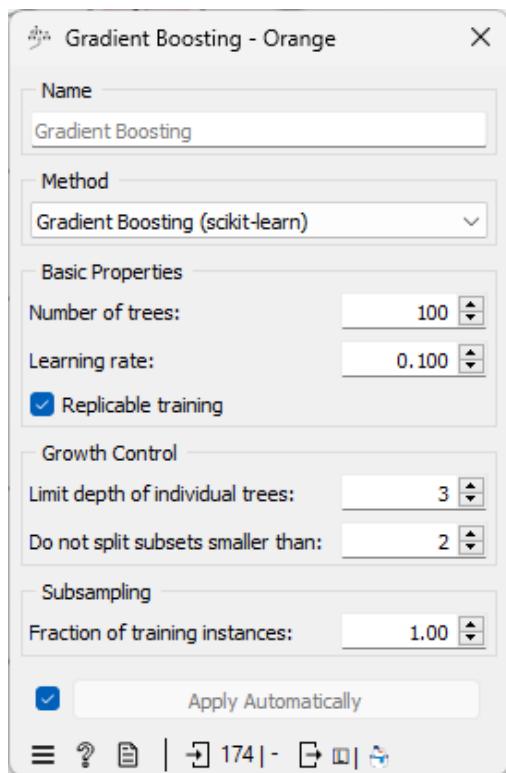
## SVM:



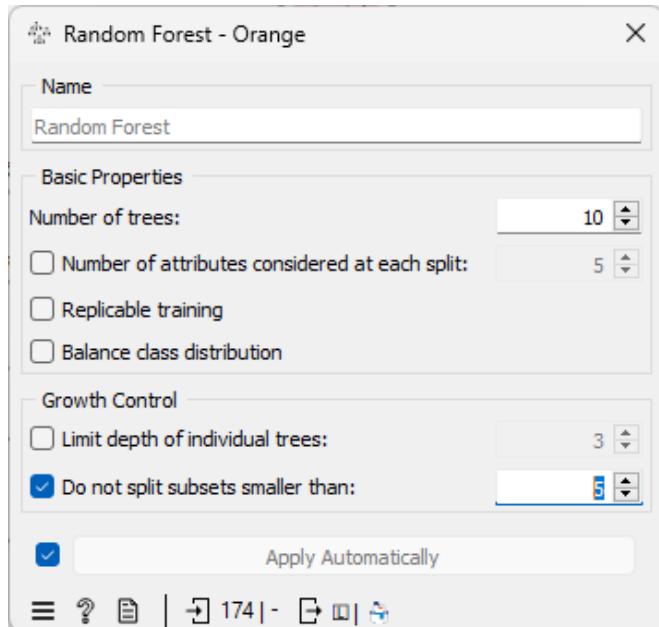
## Decision Tree:



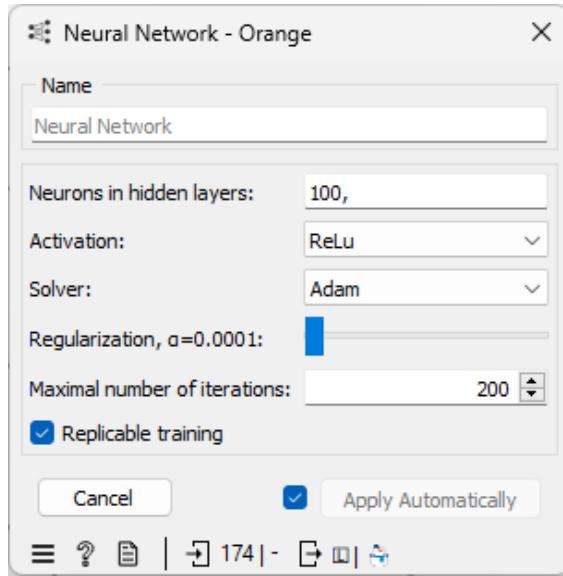
XGB:



Random Forest:



## Neural Network ReLu activation:



## Hasil Prediction:

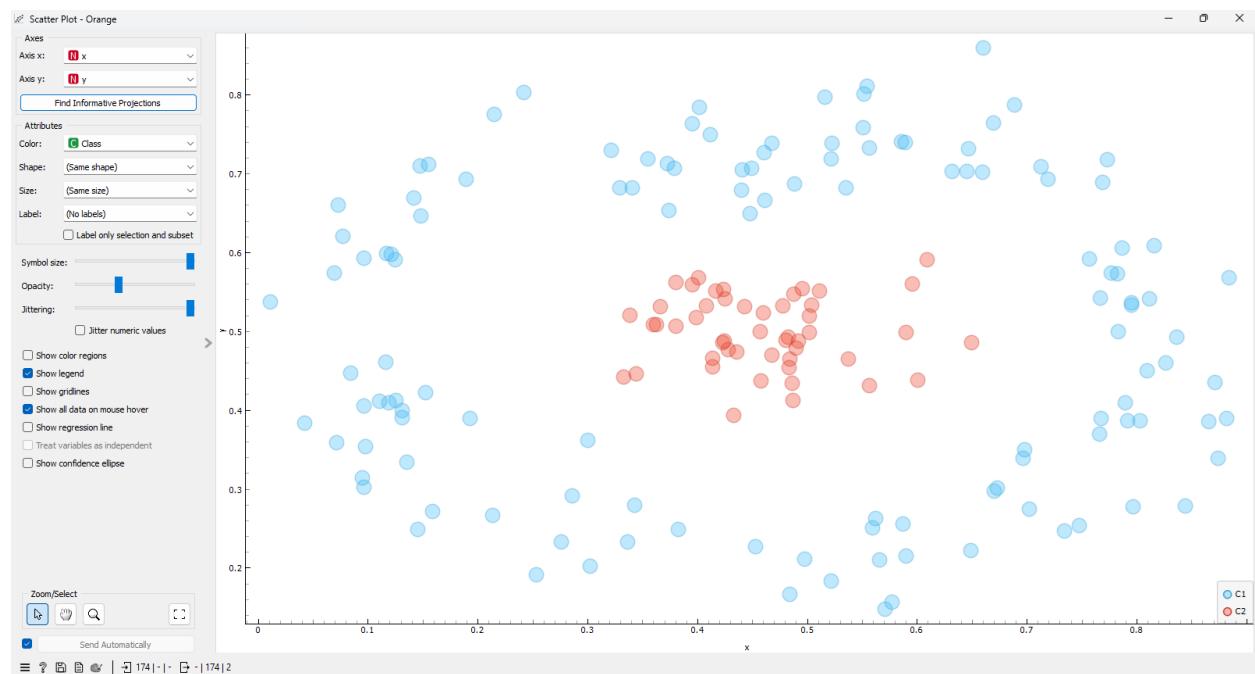
Predictions - Orange										Restore Original Order		
	Show probabilities for: Classes in data			Show classification errors								
1	SVM	error	Logistic Regression	error	Neural Network	error	Gradient Boosting	error	Tree	error	Random Forest	error
1	0.99 : 0.01 → C1	0.013	0.72 : 0.28 → C1	0.281	0.81 : 0.19 → C1	0.186	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
2	0.92 : 0.08 → C1	0.080	0.72 : 0.28 → C1	0.281	0.76 : 0.24 → C1	0.244	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	0.90 : 0.10 → C1	0.100
3	0.99 : 0.01 → C1	0.012	0.72 : 0.28 → C1	0.281	0.81 : 0.19 → C1	0.187	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
4	0.99 : 0.01 → C1	0.006	0.72 : 0.28 → C1	0.277	0.86 : 0.14 → C1	0.138	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
5	1.00 : 0.00 → C1	0.004	0.72 : 0.28 → C1	0.279	0.86 : 0.14 → C1	0.139	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
6	0.98 : 0.02 → C1	0.016	0.72 : 0.28 → C1	0.276	0.84 : 0.16 → C1	0.162	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
7	1.00 : 0.00 → C1	0.003	0.72 : 0.28 → C1	0.276	0.89 : 0.13 → C1	0.105	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
8	0.99 : 0.02 → C1	0.023	0.73 : 0.27 → C1	0.275	0.82 : 0.17 → C1	0.167	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
9	1.00 : 0.00 → C1	0.001	0.72 : 0.28 → C1	0.276	0.91 : 0.09 → C1	0.092	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
10	0.99 : 0.02 → C1	0.024	0.72 : 0.28 → C1	0.278	0.82 : 0.18 → C1	0.181	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
11	1.00 : 0.00 → C1	0.005	0.72 : 0.28 → C1	0.277	0.87 : 0.13 → C1	0.133	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
12	0.94 : 0.06 → C1	0.060	0.72 : 0.28 → C1	0.278	0.79 : 0.21 → C1	0.206	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
13	0.85 : 0.15 → C1	0.153	0.72 : 0.28 → C1	0.279	0.75 : 0.25 → C1	0.246	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
14	1.00 : 0.00 → C1	0.003	0.73 : 0.27 → C1	0.274	0.89 : 0.11 → C1	0.113	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
15	0.99 : 0.01 → C1	0.006	0.73 : 0.27 → C1	0.271	0.88 : 0.12 → C1	0.119	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	0.98 : 0.02 → C1	0.017
16	0.99 : 0.01 → C1	0.006	0.73 : 0.27 → C1	0.271	0.88 : 0.12 → C1	0.123	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	0.98 : 0.02 → C1	0.017
17	0.98 : 0.02 → C1	0.024	0.73 : 0.27 → C1	0.270	0.88 : 0.12 → C1	0.119	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
18	0.98 : 0.02 → C1	0.019	0.73 : 0.27 → C1	0.270	0.89 : 0.11 → C1	0.110	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
19	1.00 : 0.00 → C1	0.002	0.73 : 0.27 → C1	0.267	0.94 : 0.08 → C1	0.064	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
20	0.99 : 0.01 → C1	0.009	0.73 : 0.27 → C1	0.269	0.91 : 0.09 → C1	0.093	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
21	1.00 : 0.00 → C1	0.005	0.73 : 0.27 → C1	0.268	0.92 : 0.08 → C1	0.075	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
22	0.98 : 0.02 → C1	0.016	0.73 : 0.27 → C1	0.270	0.89 : 0.11 → C1	0.106	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	1.00 : 0.00 → C1	0.000
23	0.98 : 0.02 → C1	0.024	0.73 : 0.27 → C1	0.272	0.87 : 0.13 → C1	0.132	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	0.90 : 0.10 → C1	0.100
24	0.99 : 0.01 → C1	0.009	0.73 : 0.27 → C1	0.270	0.90 : 0.10 → C1	0.095	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	0.99 : 0.01 → C1	0.014
25	1.00 : 0.00 → C1	0.005	0.73 : 0.27 → C1	0.270	0.91 : 0.09 → C1	0.093	1.00 : 0.00 → C1	0.001	1.00 : 0.00 → C1	0.000	0.90 : 0.10 → C1	0.100

Show performance scores							Target class: (Average over classes)					
Model	AUC	CA	F1	Prec	Recall	MCC						
SVM	1.000	1.000	1.000	1.000	1.000	1.000						
Logistic Regression	0.558	0.718	0.601	0.598	0.716	0.000						
Neural Network	1.000	0.914	0.908	0.923	0.914	0.787						
Gradient Boosting	1.000	1.000	1.000	1.000	1.000	1.000						
Tree	0.971	0.960	0.960	0.963	0.960	0.906						
Random Forest	1.000	0.994	0.994	0.994	0.994	0.986						

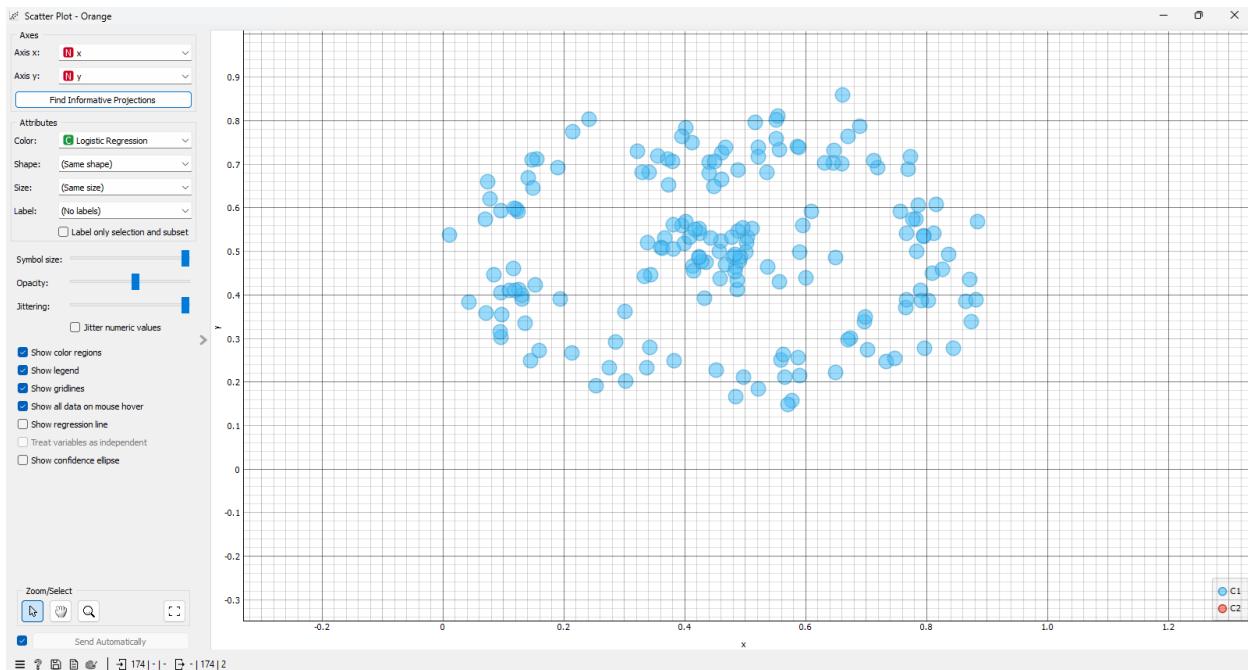
Berdasarkan hasil yang ditampilkan, **Logistic Regression** menunjukkan performa yang buruk dengan AUC hanya 0.558, akurasi 71,8%, F1 score 0.601, dan MCC 0, yang berarti model ini tidak efektif dalam membedakan kelas positif dan negatif. Precision yang sangat rendah (0.516) menunjukkan banyaknya kesalahan dalam mengklasifikasikan kelas negatif sebagai positif, meskipun recall yang lebih tinggi (0.718) menunjukkan model lebih baik dalam mendeteksi kelas positif.

Sebaliknya, **Tree**, **Random Forest**, **Gradient Boosting**, **Neural Network**, dan **SVM** menunjukkan performa yang sangat baik dengan AUC, akurasi, F1 score, precision, recall, dan MCC yang hampir sempurna (semua mencapai 1.000). Model-model ini mampu mengklasifikasikan hampir semua data dengan benar dan seimbang dalam mendeteksi kedua kelas, menjadikannya pilihan terbaik untuk tugas klasifikasi ini. Oleh karena itu, model-model seperti **Gradient Boosting** dan **SVM** lebih disarankan, sementara **Logistic Regression** perlu ditingkatkan atau diganti dengan model yang lebih kuat.

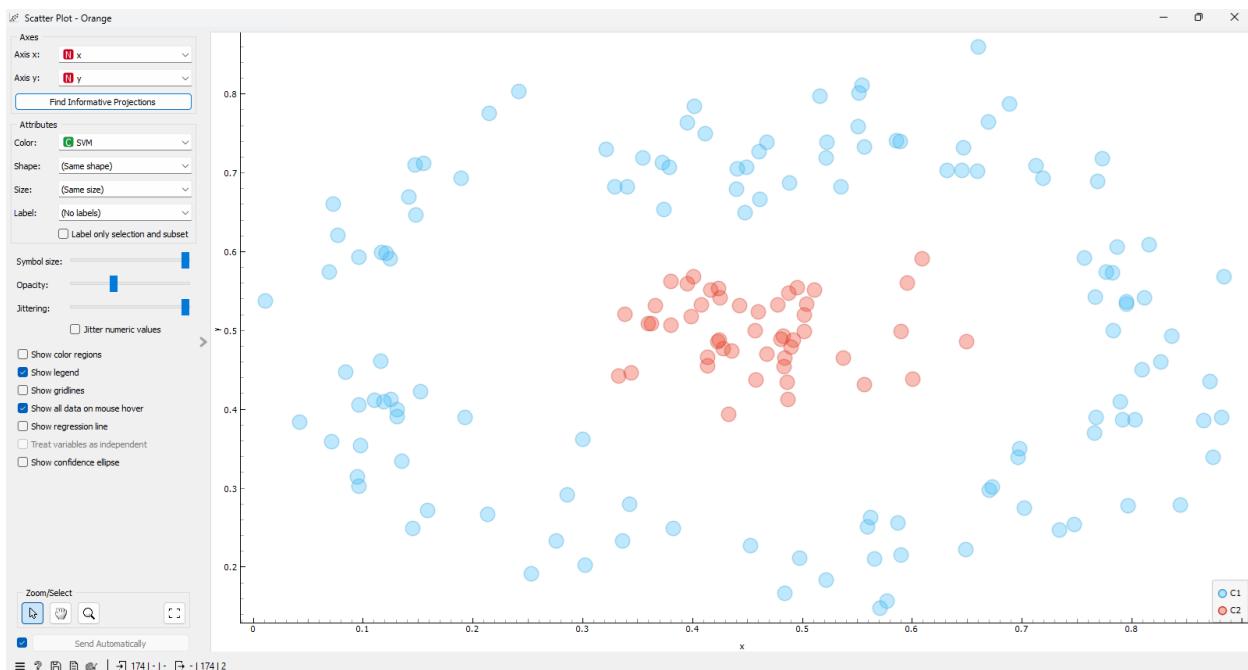
### Scatter Plot



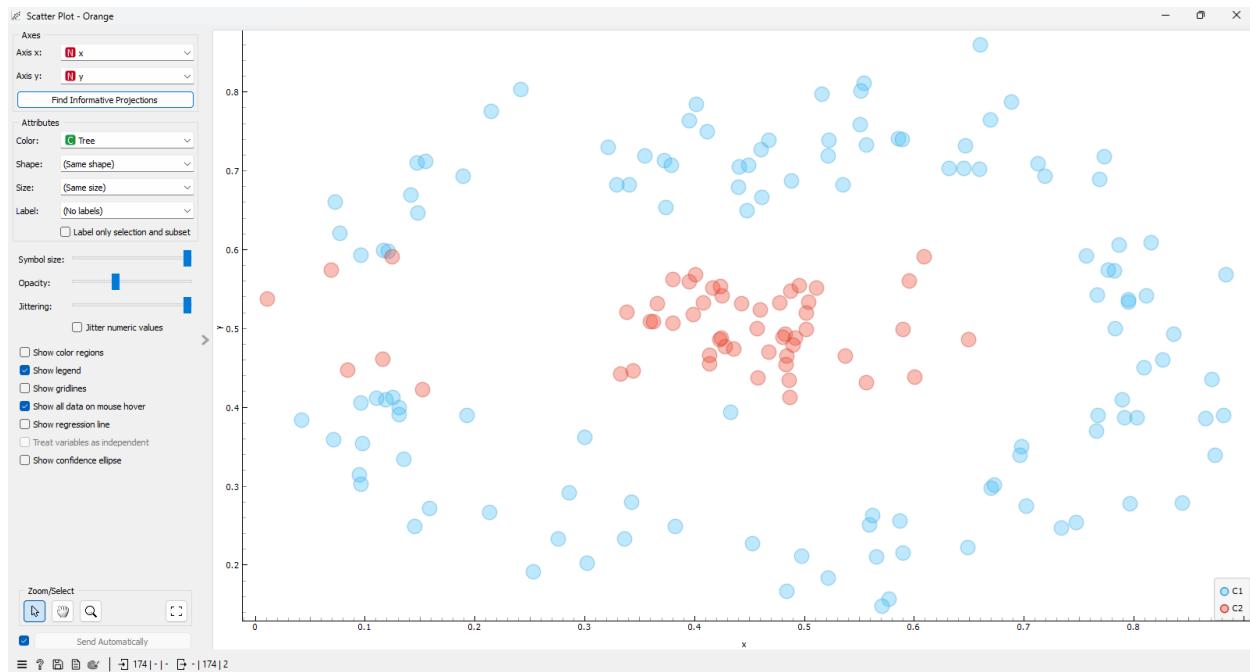
## Logistic Regression



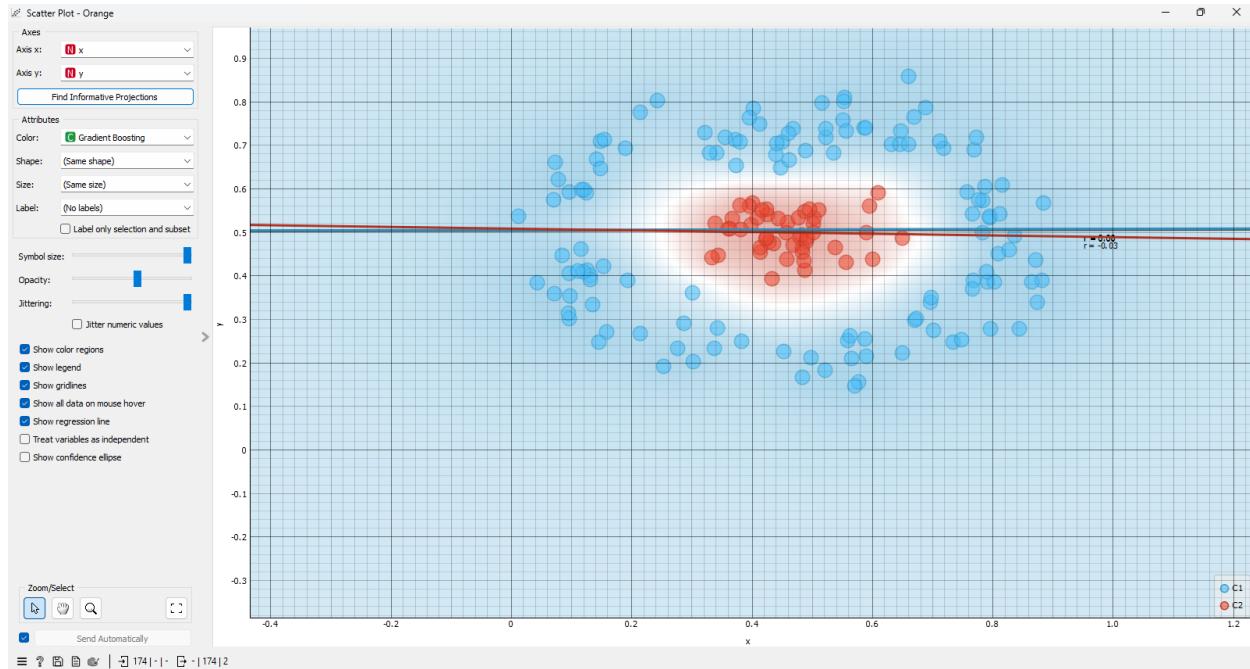
## SVM



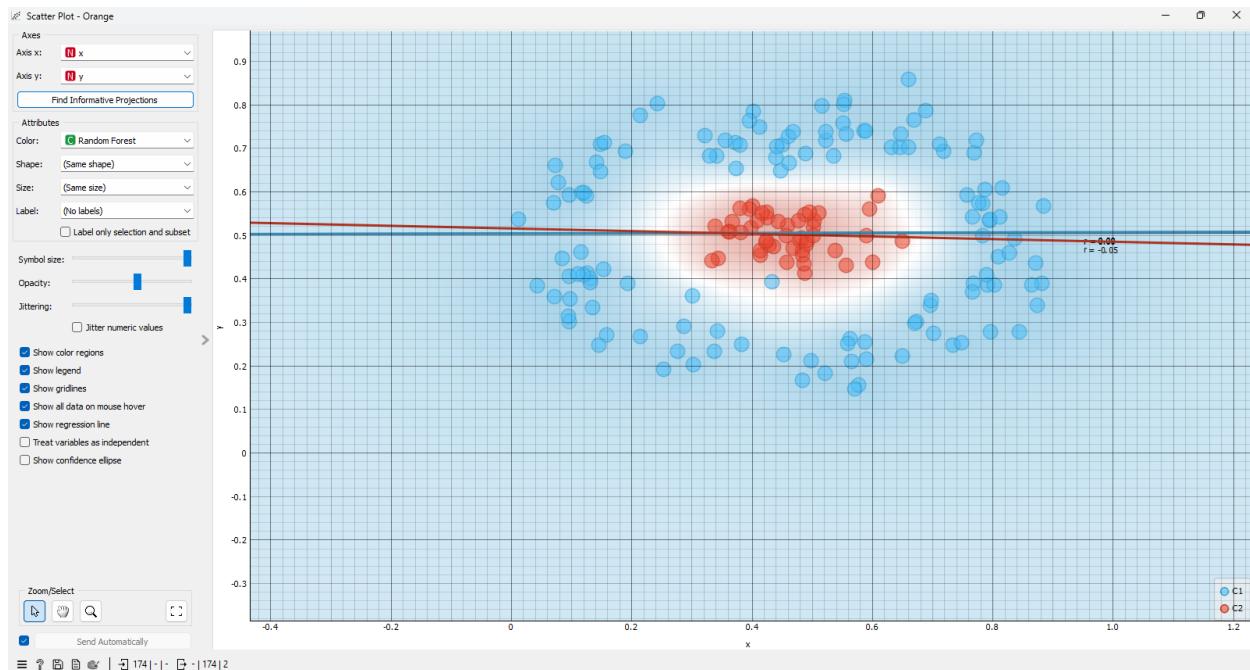
## Decision Tree



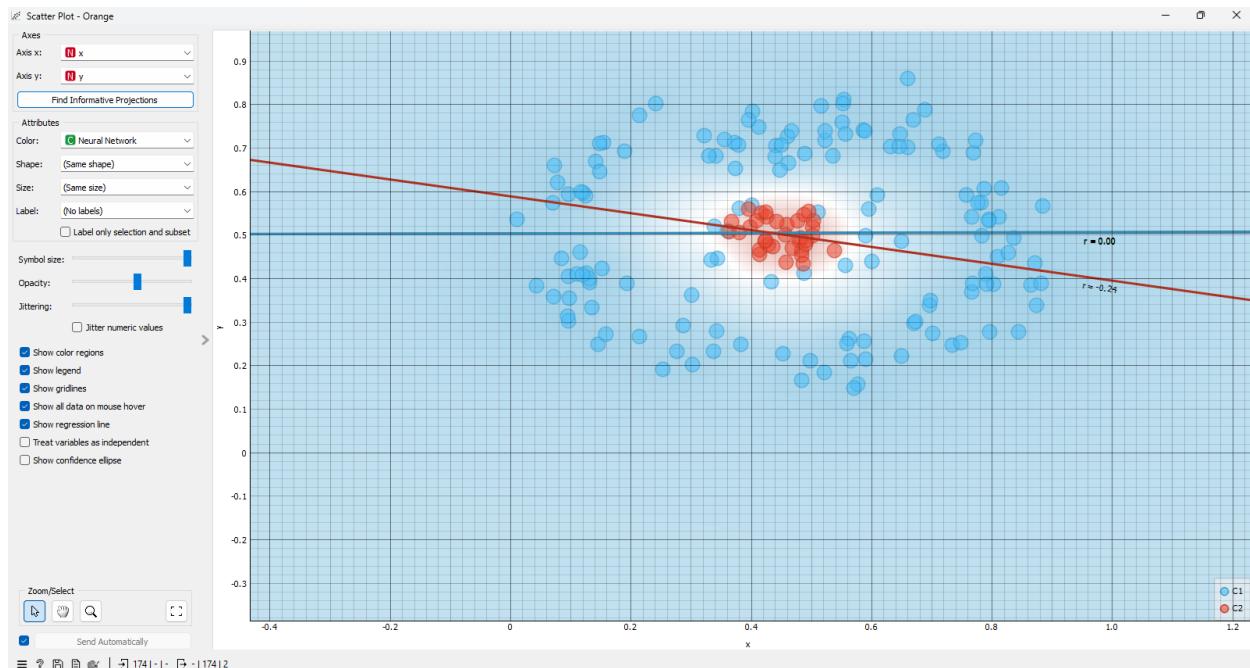
## XGB



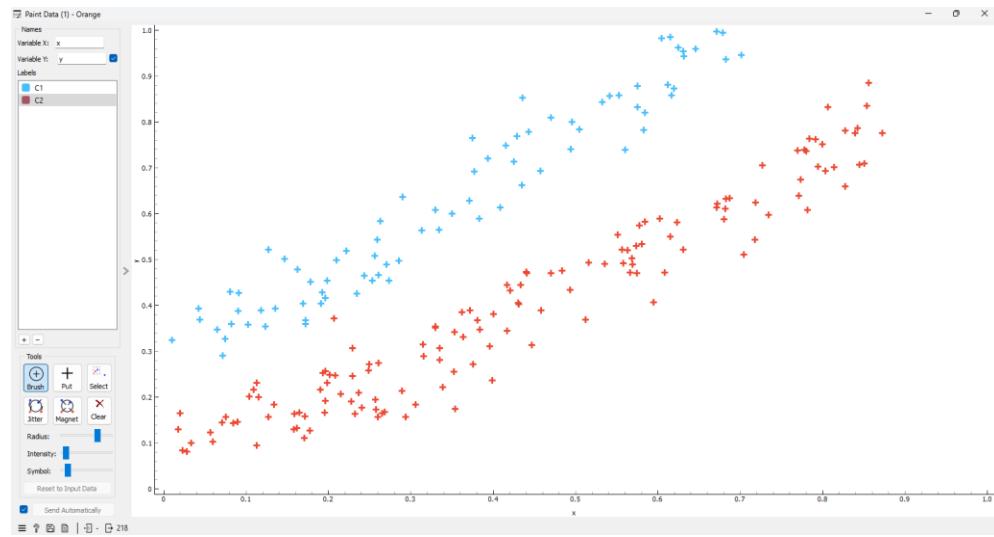
## Random Forest



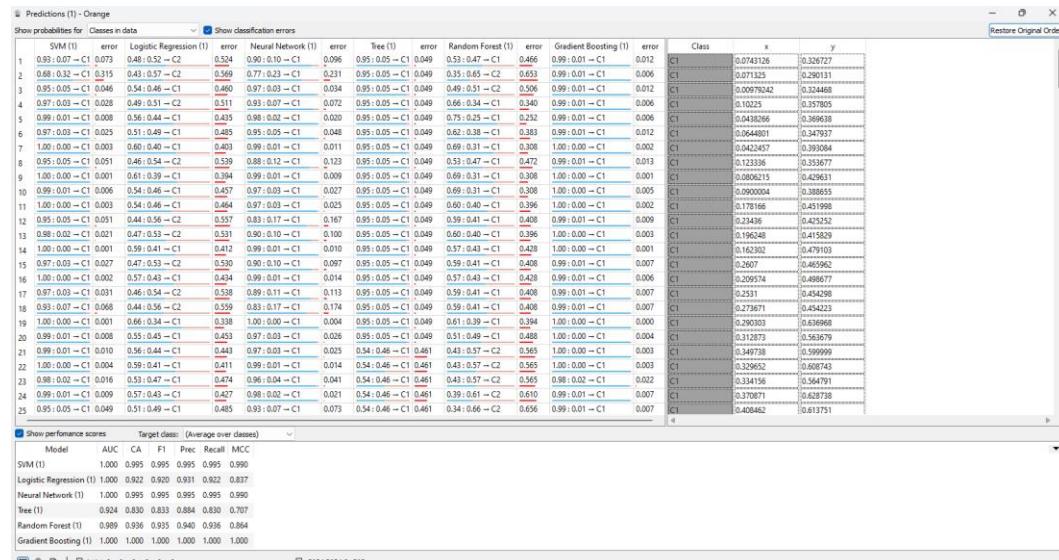
## Neural Network ReLu activation



## Paint data Pola 2



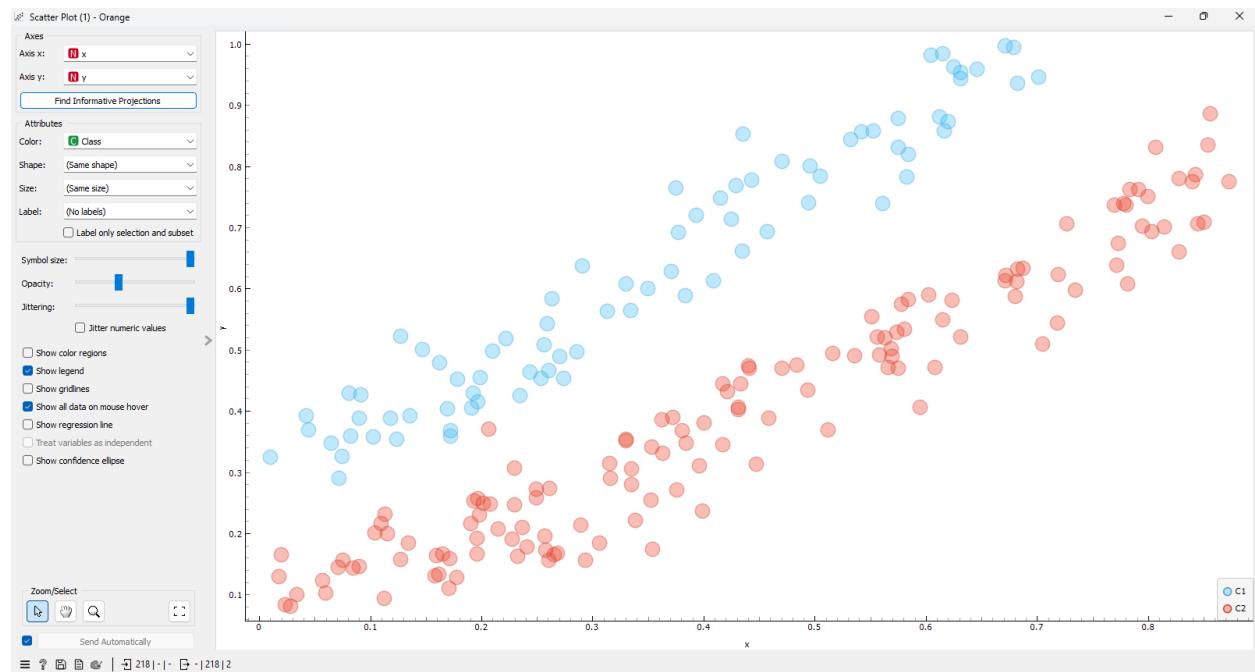
## Hasil Prediction



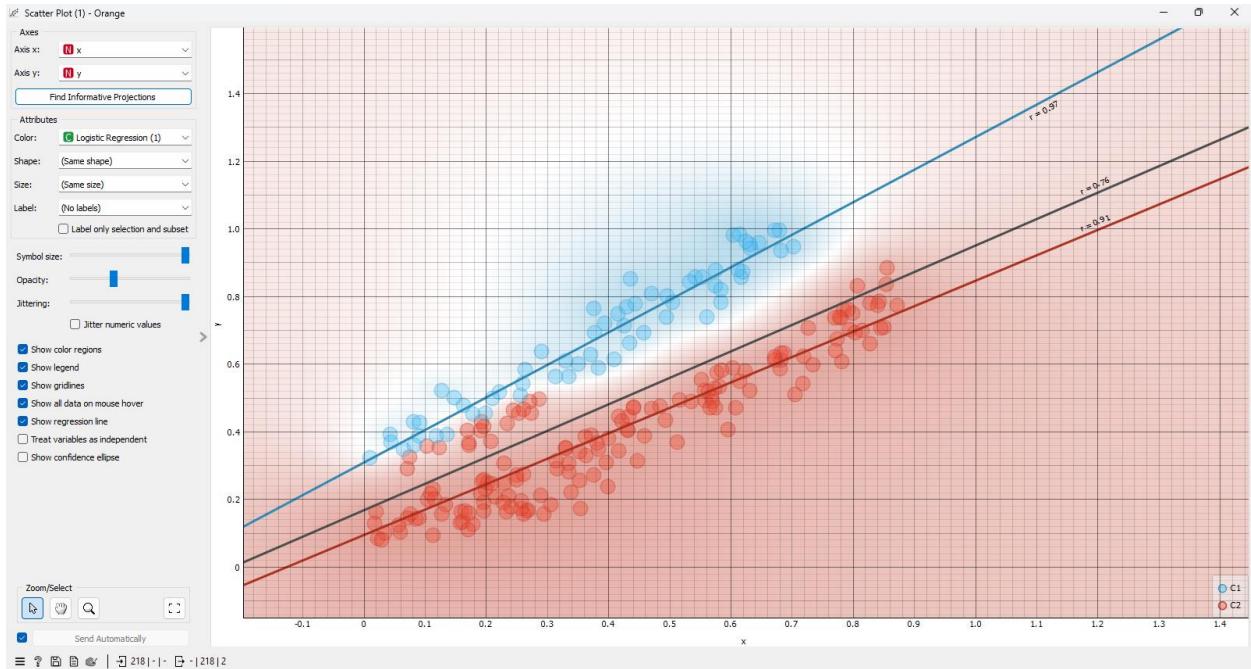
Show performance scores		Target class: (Average over classes)					
Model	AUC	CA	F1	Prec	Recall	MCC	
SVM (1)	1.000	0.995	0.995	0.995	0.990		
Logistic Regression (1)	1.000	0.922	0.920	0.911	0.922	0.837	
Neural Network (1)	1.000	0.995	0.995	0.995	0.995	0.995	
Tree (1)	0.924	0.830	0.833	0.884	0.830	0.707	
Random Forest (1)	0.989	0.936	0.935	0.940	0.936	0.864	
Gradient Boosting (1)	1.000	1.000	1.000	1.000	1.000	1.000	

Berdasarkan hasil evaluasi model yang ditampilkan, dapat dilihat bahwa sebagian besar model menunjukkan kinerja yang sangat baik dengan nilai AUC, CA, F1, Precision, Recall, dan MCC mendekati 1.0, yang menandakan bahwa model tersebut berhasil mengklasifikasikan data dengan sangat akurat. **Logistic Regression, Neural Network, Gradient Boosting, Random Forest, dan SVM** semuanya hampir mencapai skor sempurna (1.000) untuk AUC, CA, F1, Precision, dan Recall, menunjukkan bahwa model-model ini dapat memprediksi kelas dengan tingkat kesalahan yang sangat rendah. **Tree model** sedikit lebih rendah performanya dengan AUC 0.924, CA 0.830, dan MCC 0.707, yang masih tergolong sangat baik meskipun ada sedikit penurunan dibandingkan model lainnya.

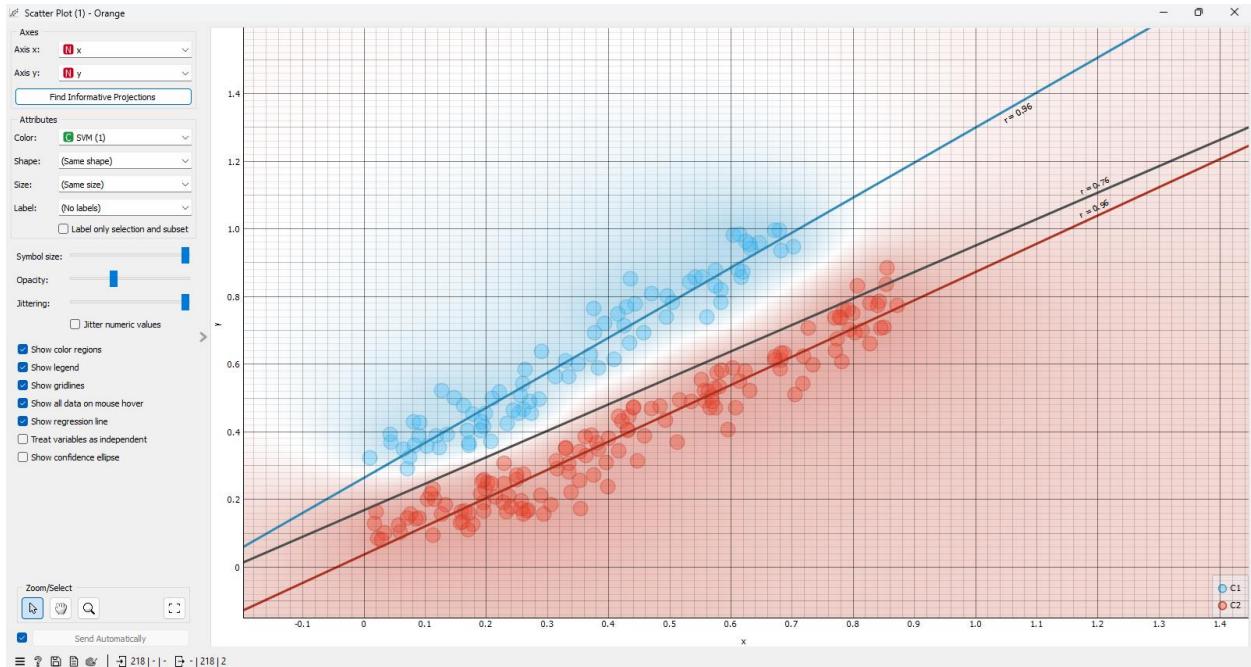
## Scatter Plot



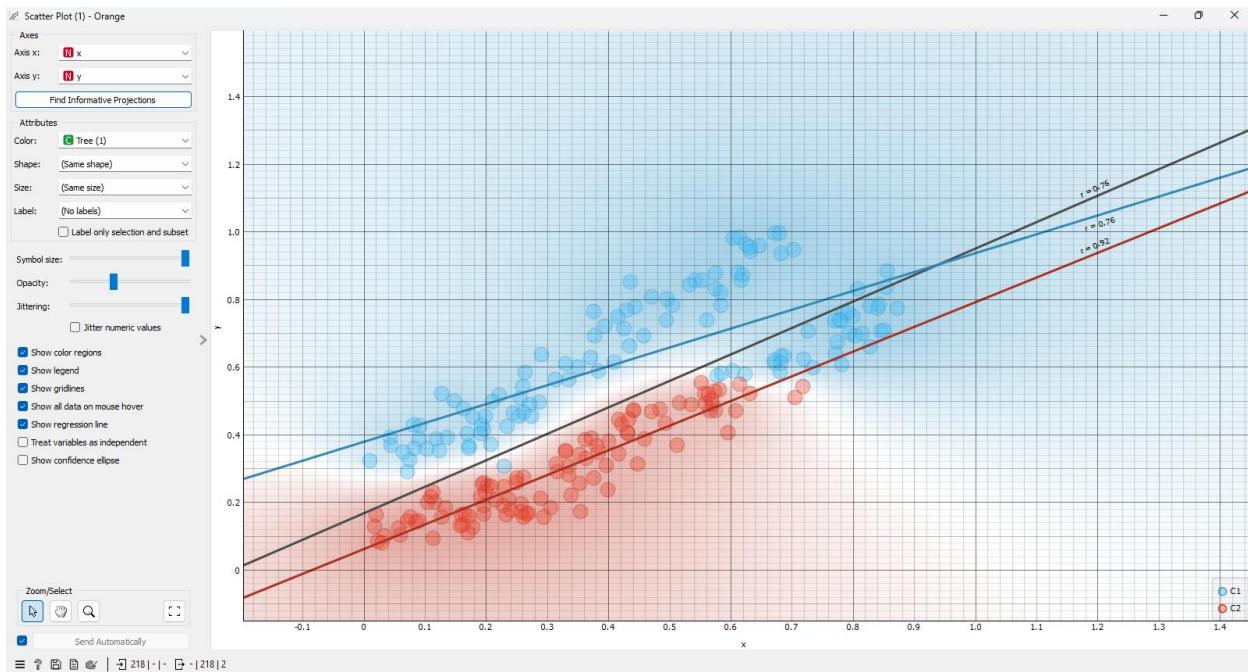
## Logistic Regression



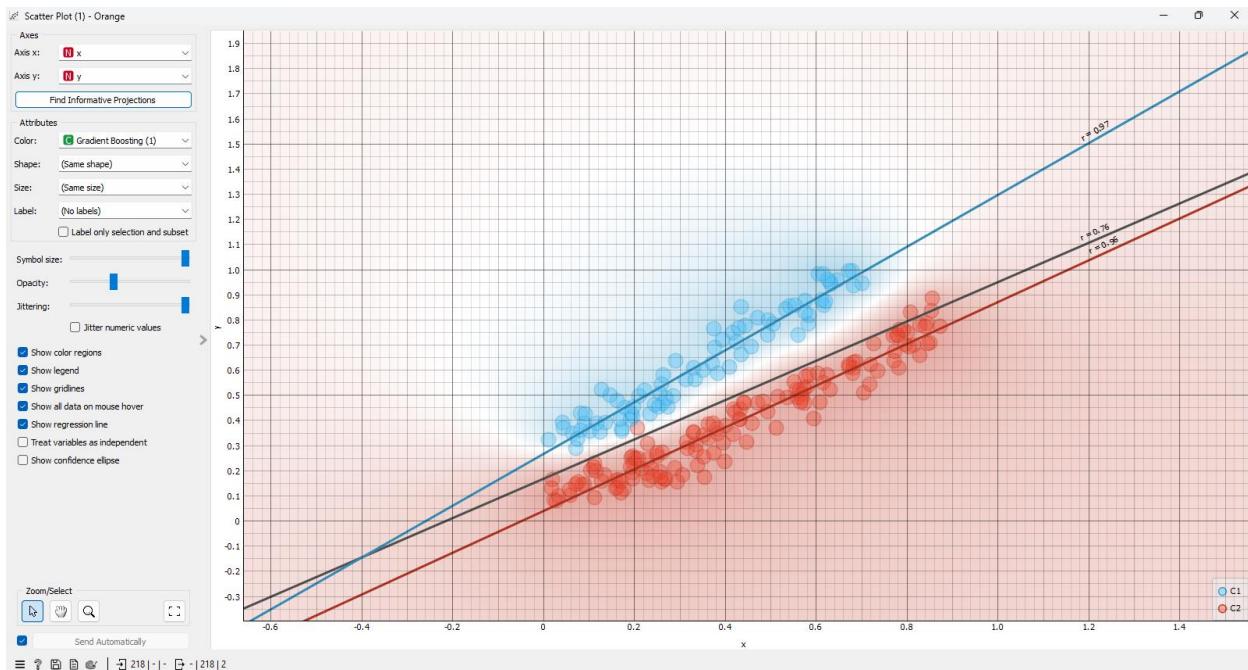
## SVM



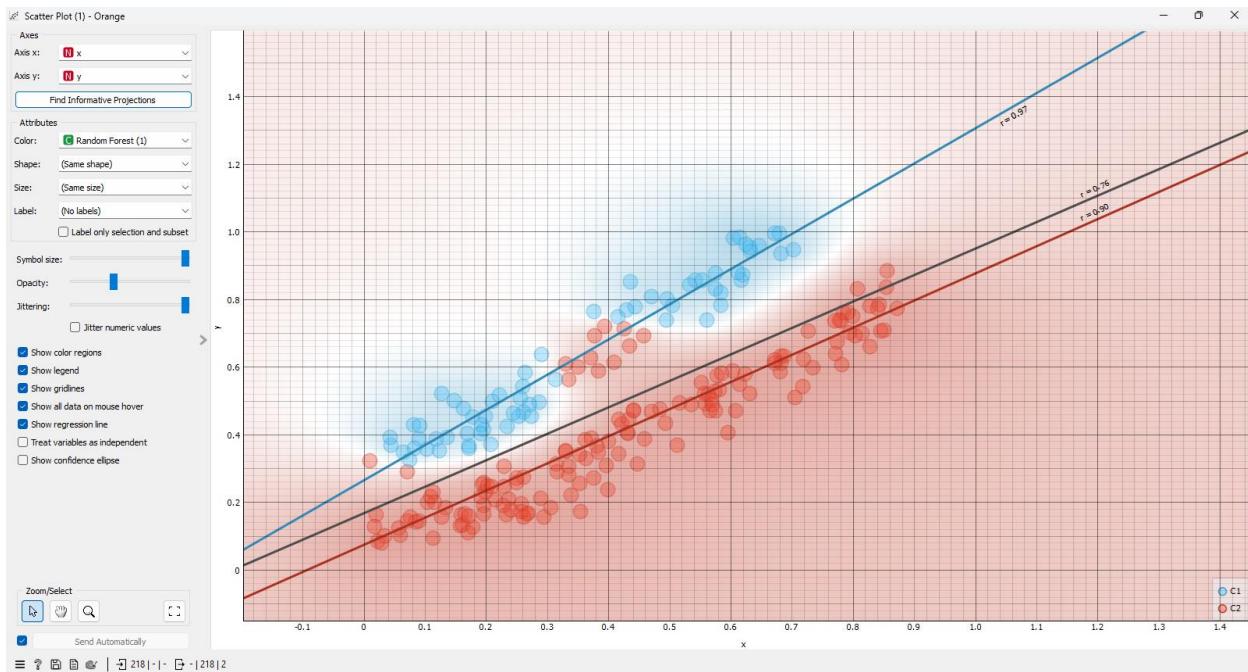
## Decision Tree



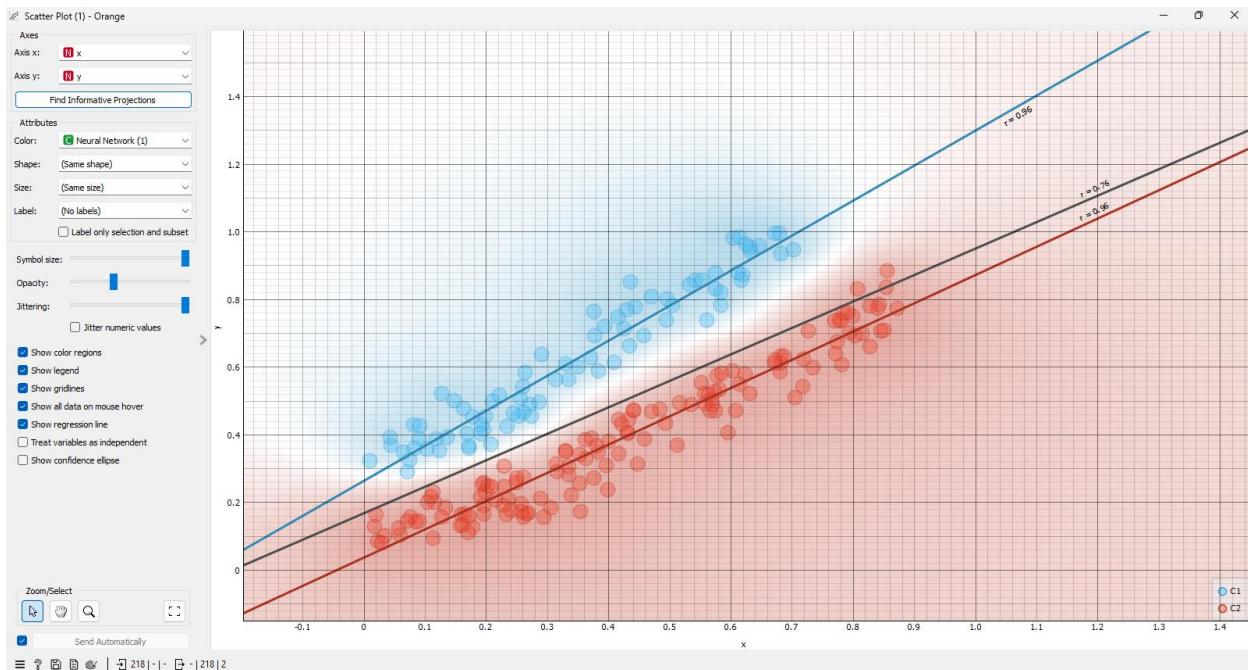
## XGB



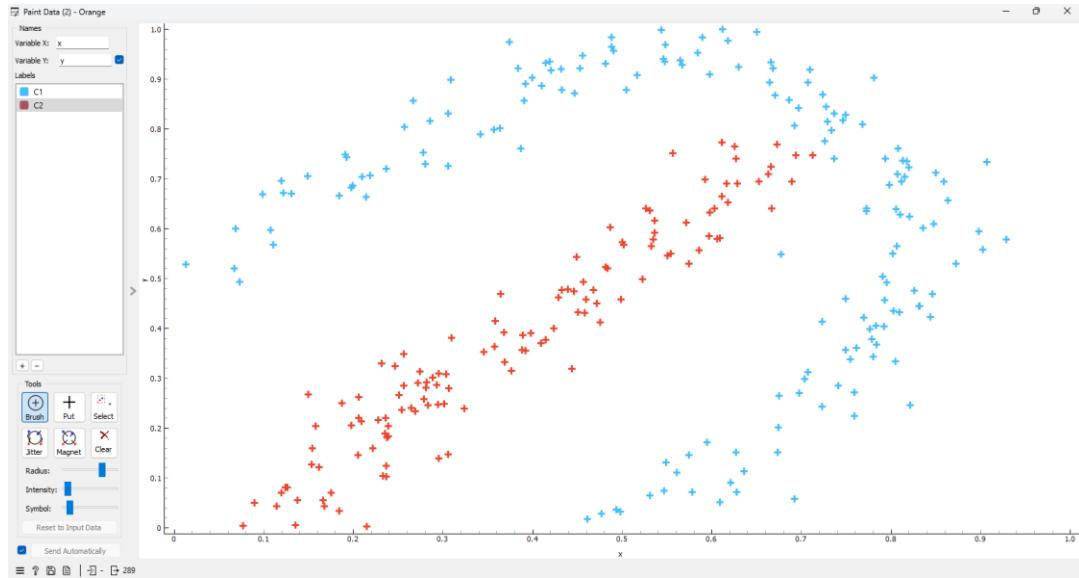
## Random Forest



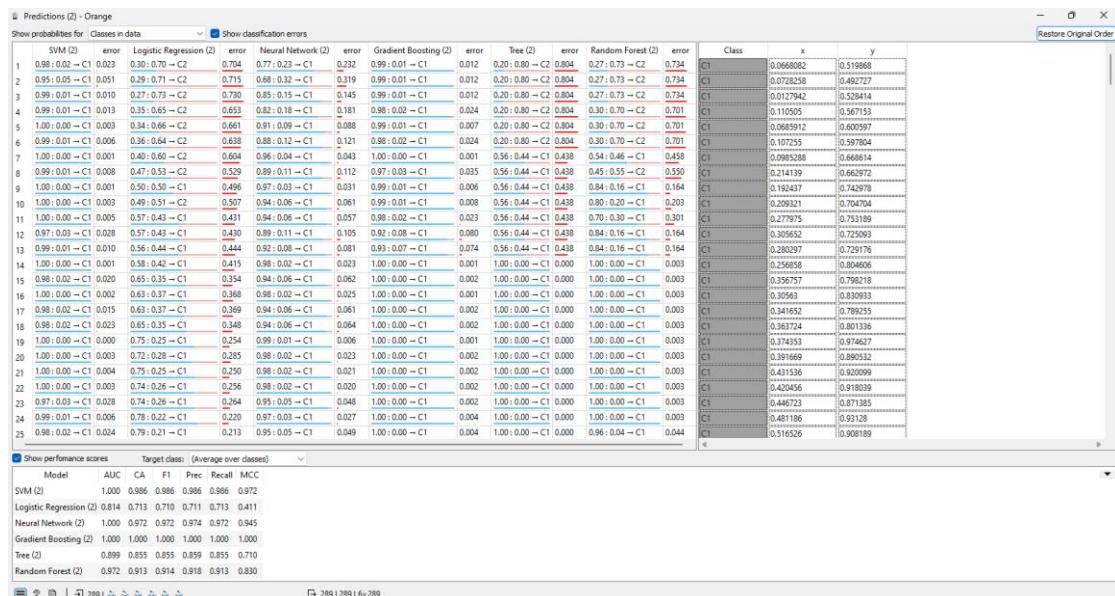
## Neural Network ReLu activation



### Paint data pola ke 3



### Hasil Prediction



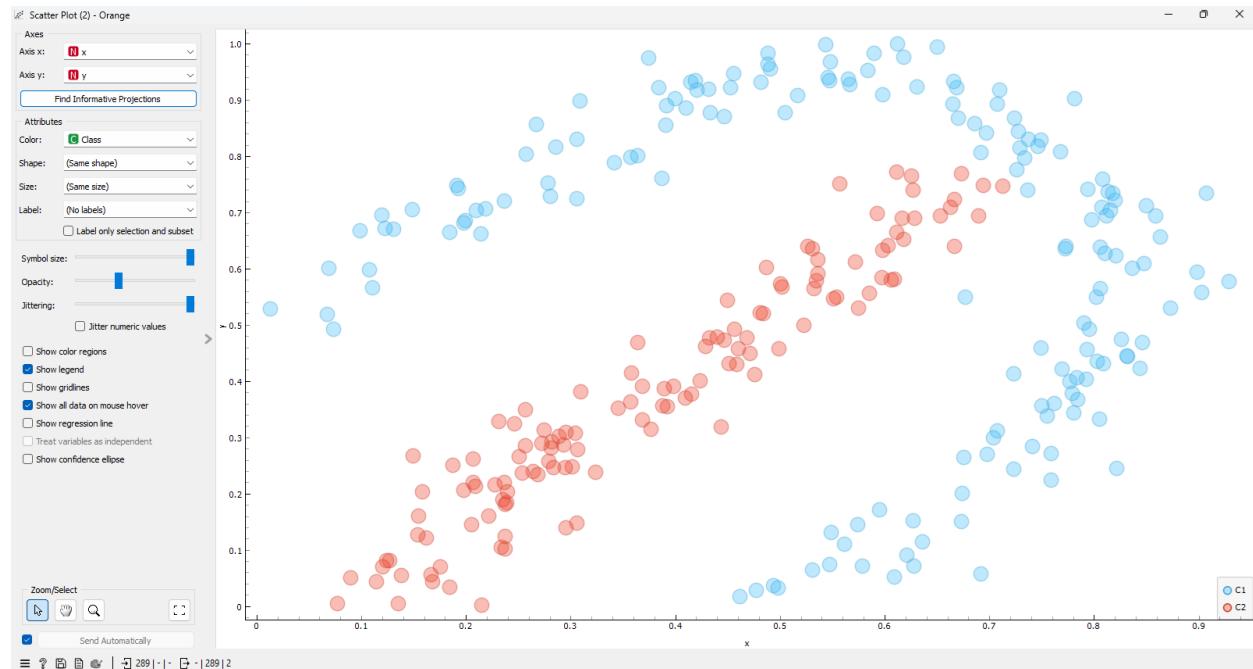
Show performance scores Target class: | (Average over classes) |

Model	AUC	CA	F1	Prec	Recall	MCC
SVM (2)	1.000	0.986	0.986	0.986	0.972	
Logistic Regression (2)	0.814	0.713	0.710	0.711	0.713	0.411
Neural Network (2)	1.000	0.972	0.972	0.974	0.945	
Gradient Boosting (2)	1.000	1.000	1.000	1.000	1.000	
Tree (2)	0.899	0.855	0.855	0.859	0.855	0.710
Random Forest (2)	0.972	0.913	0.914	0.918	0.913	0.830

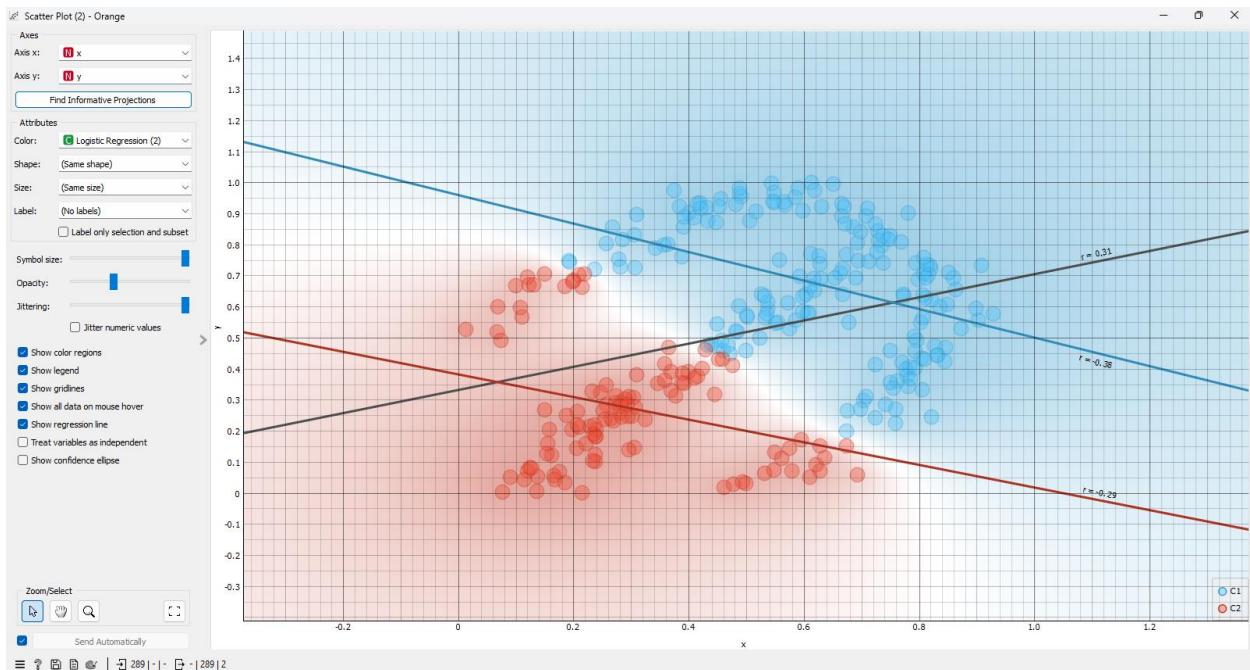
Hasil evaluasi model menunjukkan perbedaan kinerja yang signifikan antar model dalam hal AUC, CA, F1, Precision, Recall, dan MCC. Model **Gradient Boosting, Random Forest, dan SVM** memiliki performa yang sangat baik dengan nilai 1.000 di hampir semua metrik, mencerminkan kemampuannya dalam mengklasifikasikan data dengan tingkat kesalahan yang sangat rendah. **Neural Network** juga menunjukkan hasil yang sangat kuat dengan nilai AUC, CA, F1, Precision, dan Recall mendekati 1.000, meskipun sedikit lebih rendah dibandingkan dengan model Random Forest dan SVM. Sebaliknya, **Logistic Regression** mengalami penurunan yang signifikan dengan skor AUC 0.814, CA 0.713, dan F1 0.710, menunjukkan bahwa model ini kurang efektif dalam mengklasifikasikan data pada masalah ini.

Pada sisi lain, **Tree** model menunjukkan kinerja yang lebih baik dibandingkan Logistic Regression dengan nilai AUC 0.899, CA 0.855, dan F1 0.855, meskipun masih berada jauh di bawah model-model seperti **Gradient Boosting, Random Forest, dan SVM**. Meskipun Tree lebih unggul dibandingkan Logistic Regression, ia masih tidak sebanding dengan kinerja model-model ensemble dan SVM yang secara konsisten menghasilkan skor sempurna. Secara keseluruhan, model-model berbasis ensemble seperti Random Forest, Gradient Boosting, dan SVM memberikan kinerja terbaik, sementara Logistic Regression dan Tree menunjukkan performa yang lebih rendah, dengan Logistic Regression terutama membutuhkan perbaikan untuk dapat bersaing pada tingkat ini.

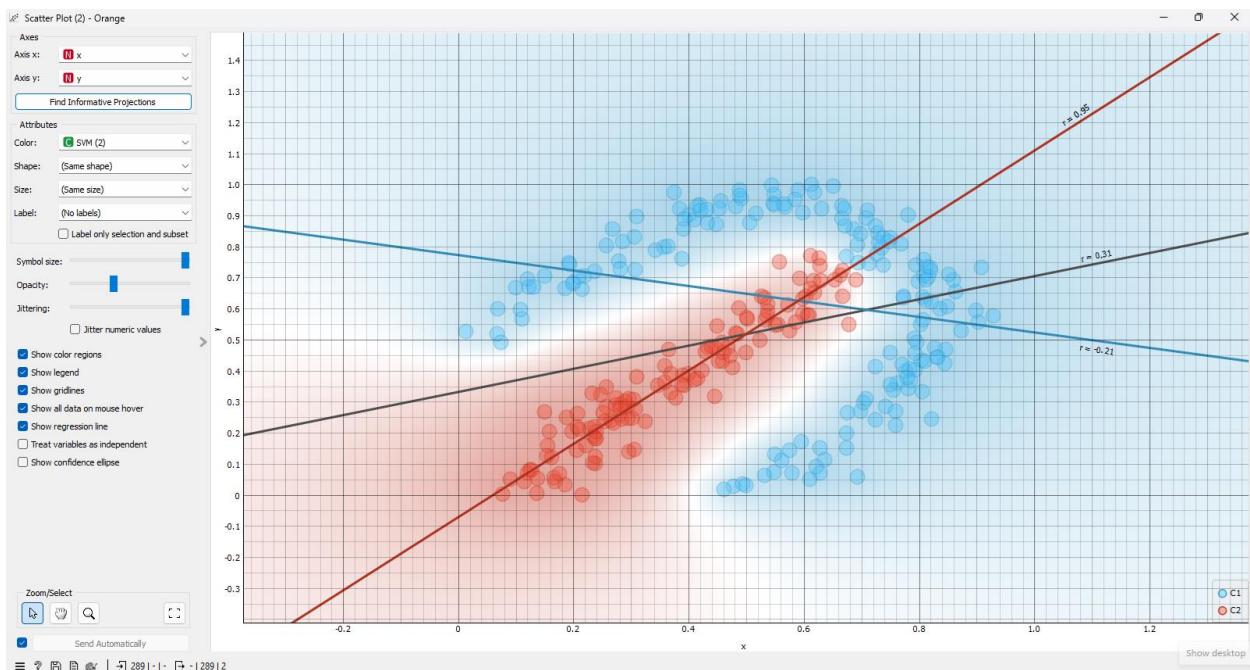
## Scatter Plot



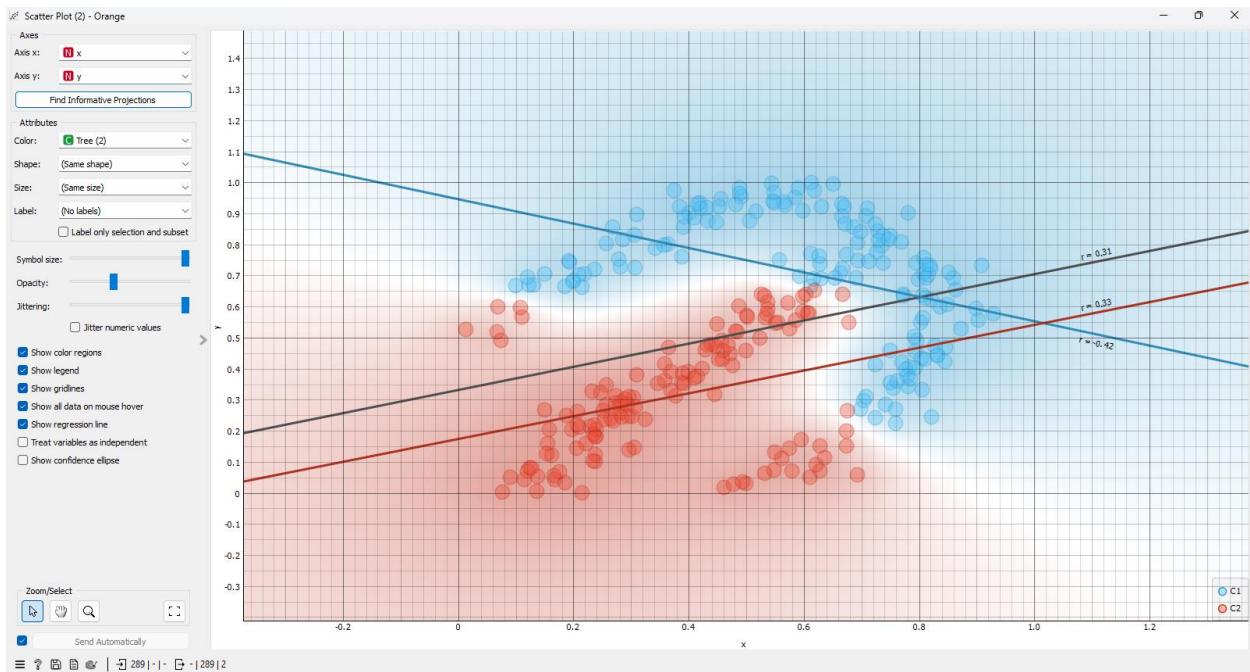
## Logistic Regression



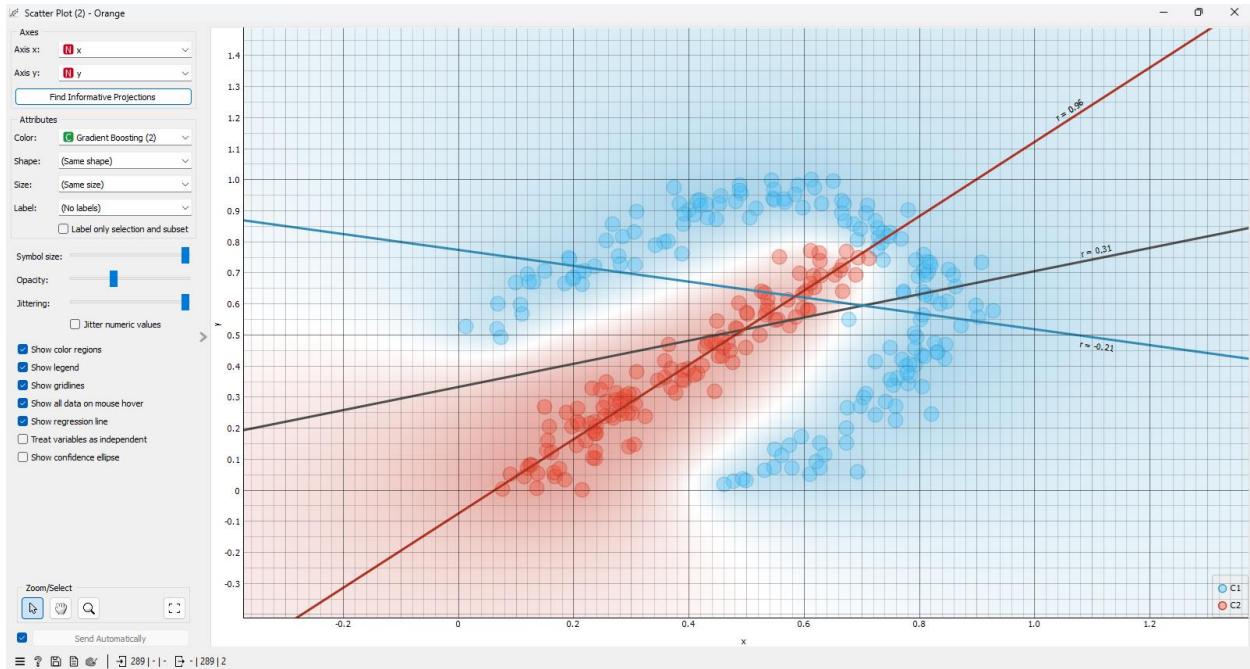
## SVM



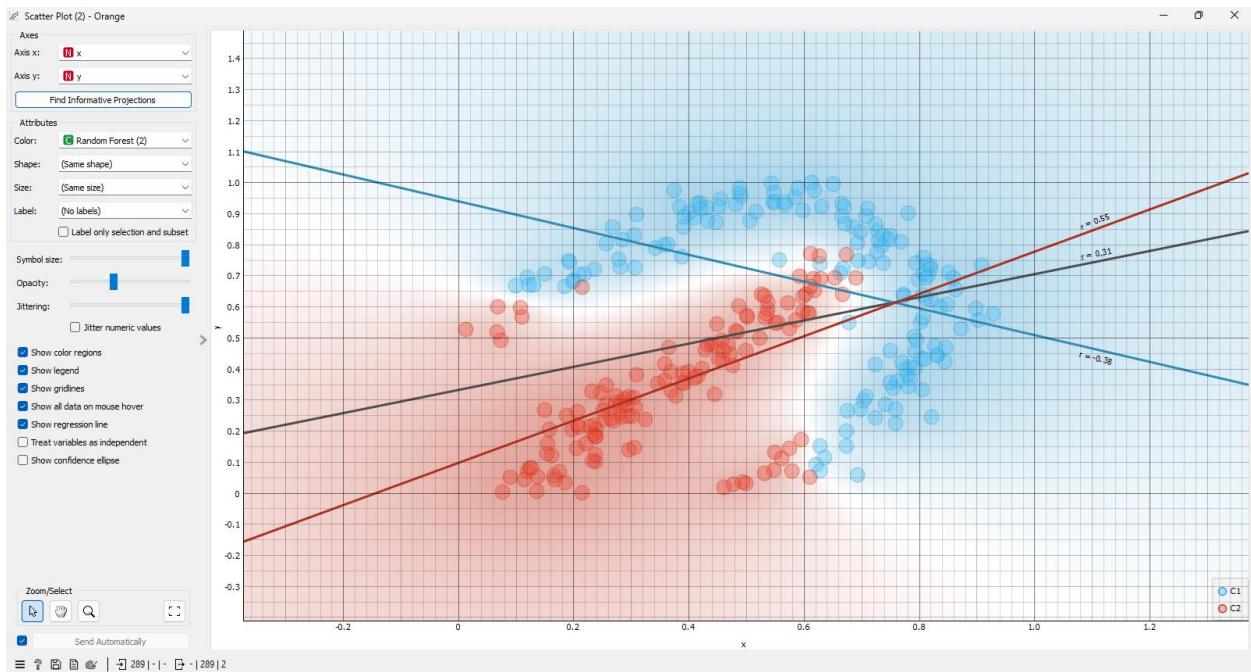
## Decision Tree



## XGB



## Random Forest



## Neural Network ReLu activation

