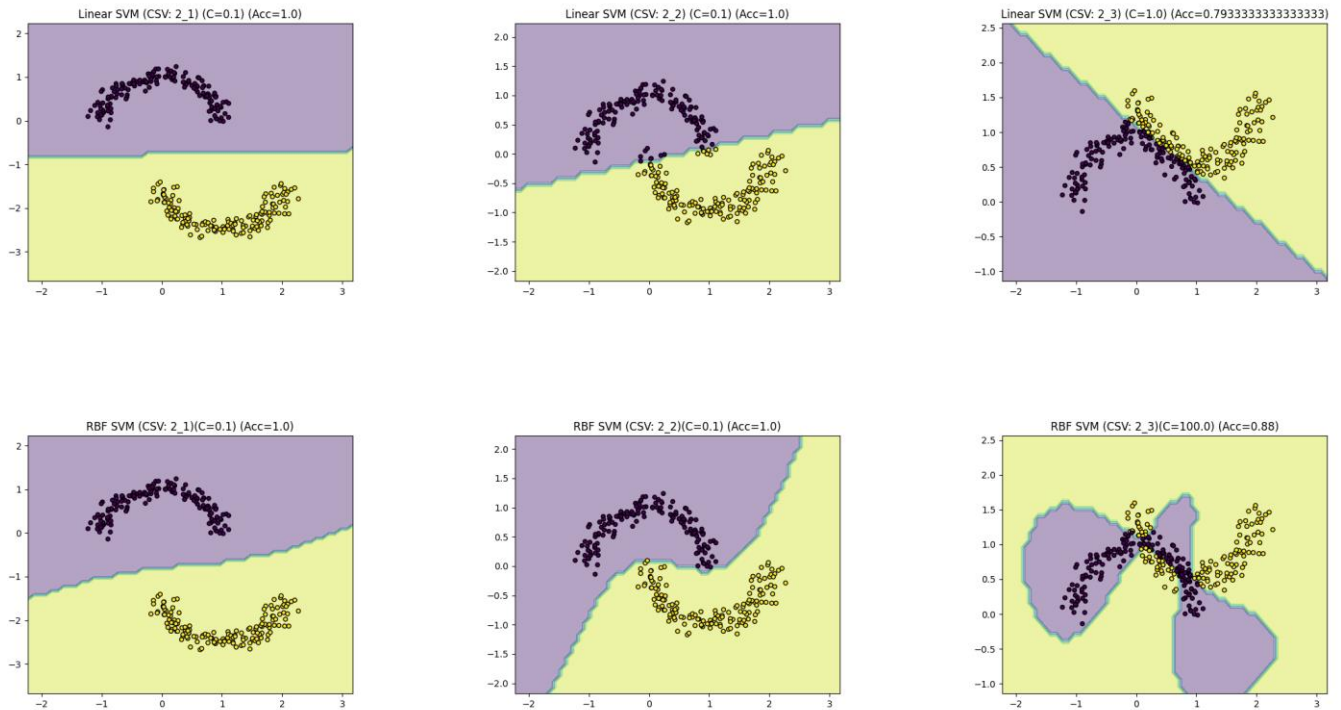
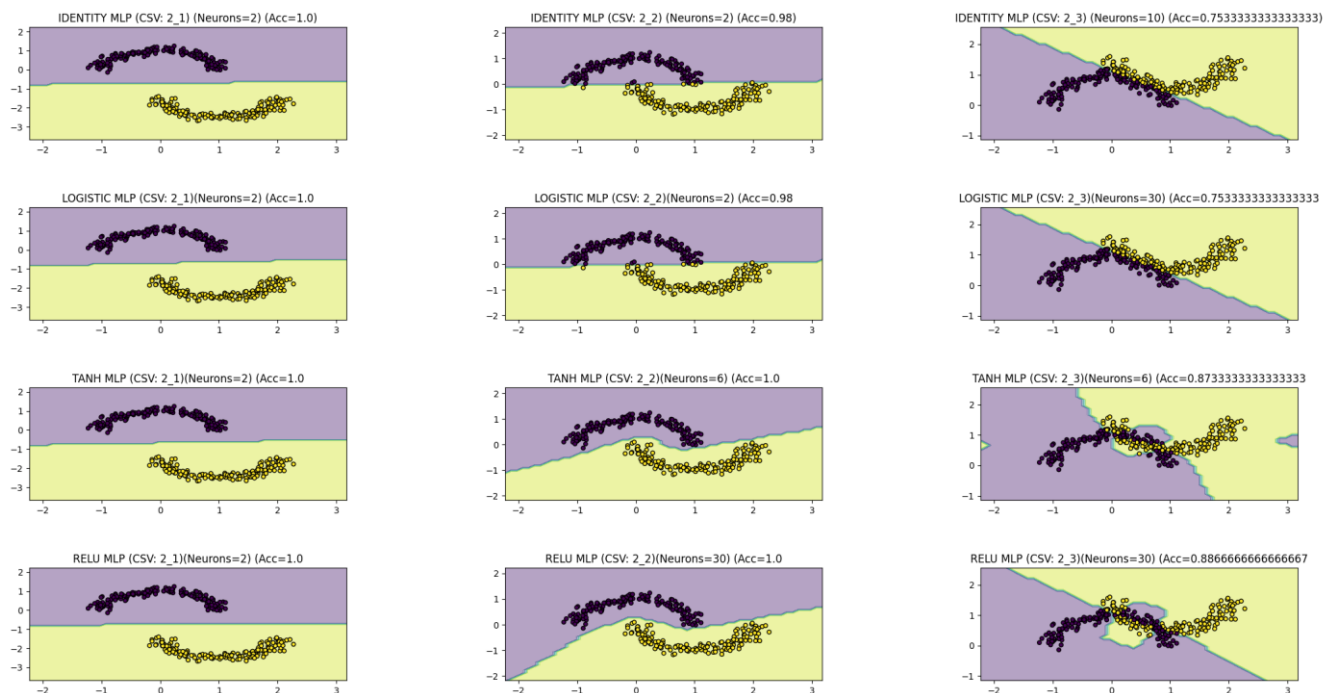


2.1 Wyniki pierwszego eksperymentu dla trzech sztucznie wygenerowanych zbiorów danych 2_1, 2_2 oraz 2_3.

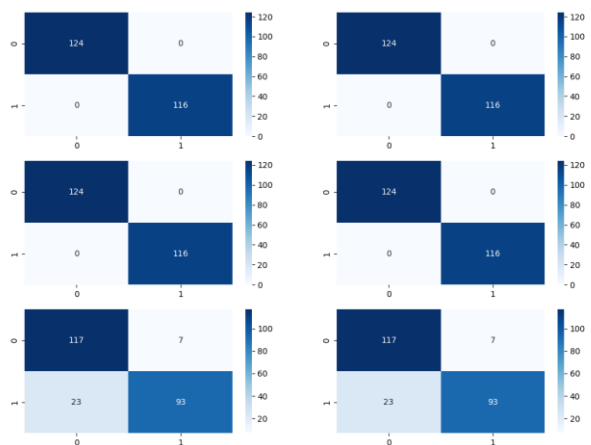
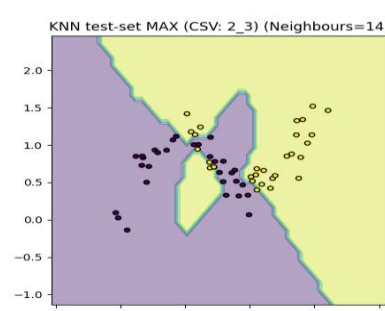
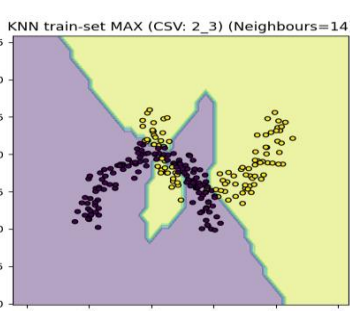
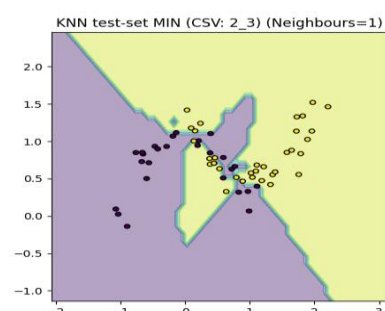
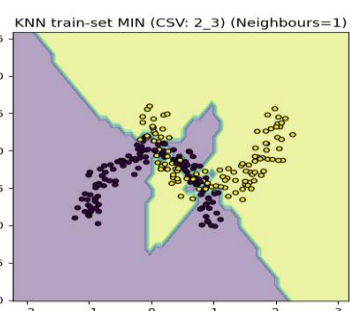
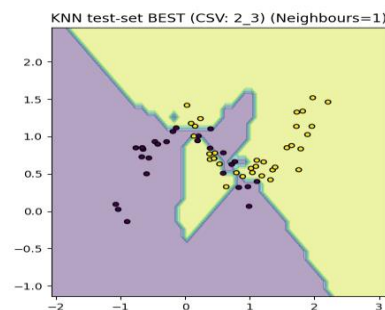
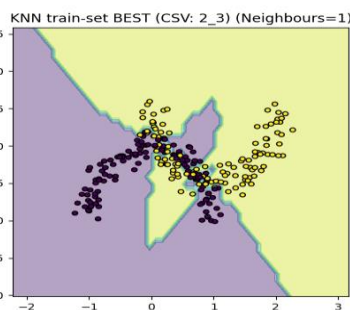
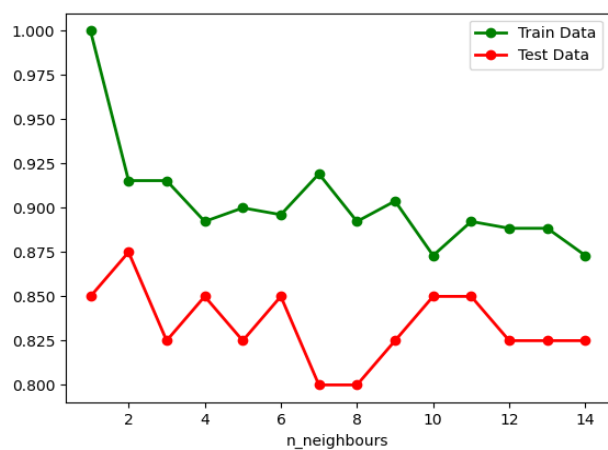
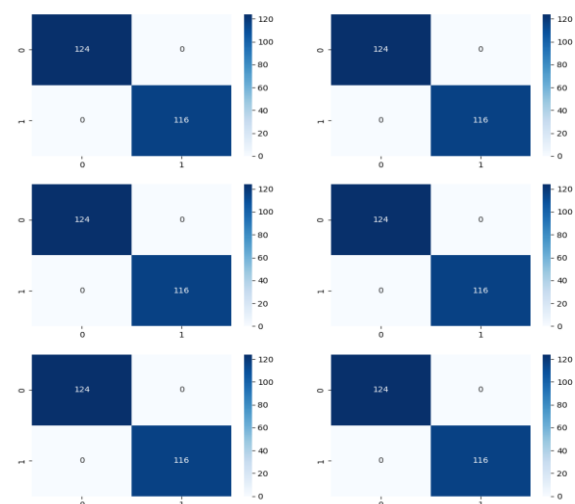
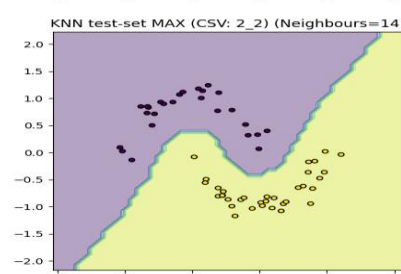
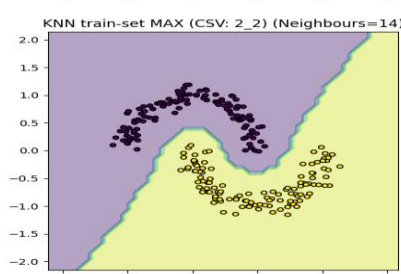
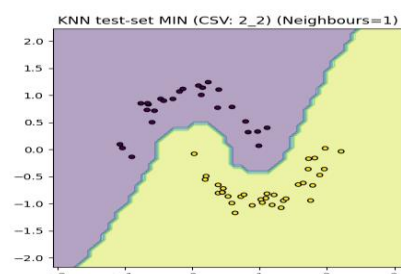
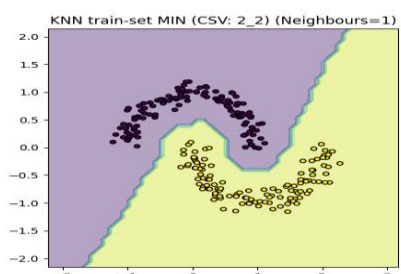
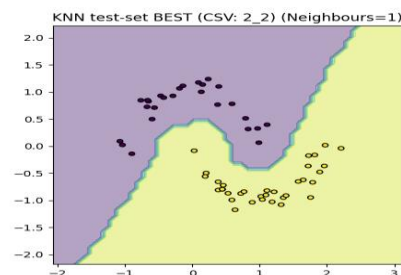
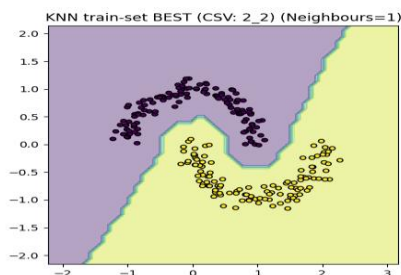
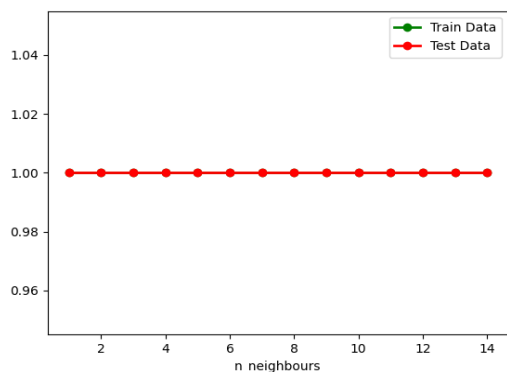
SVM



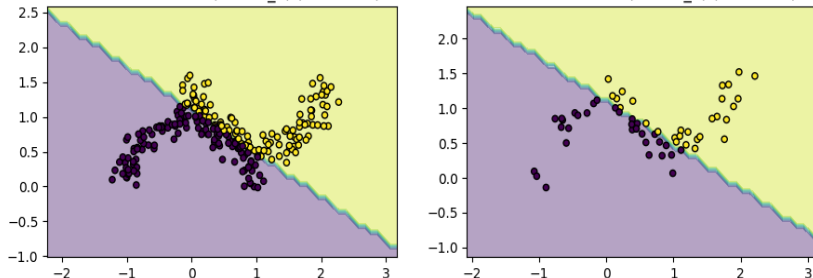
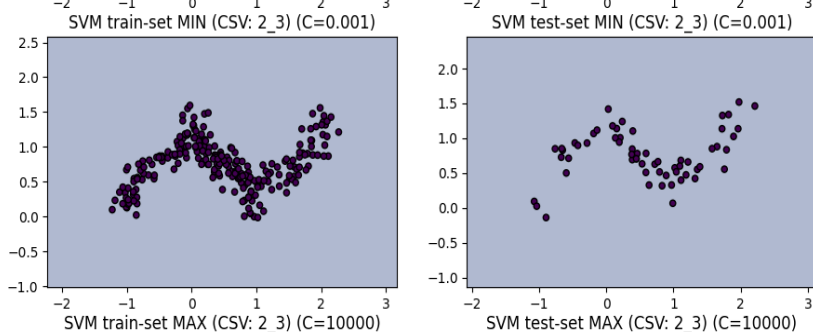
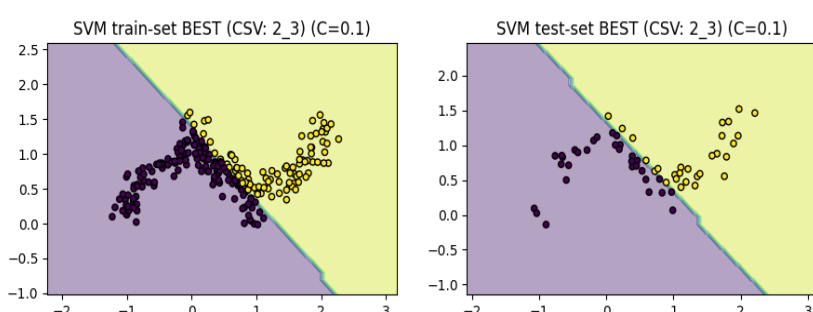
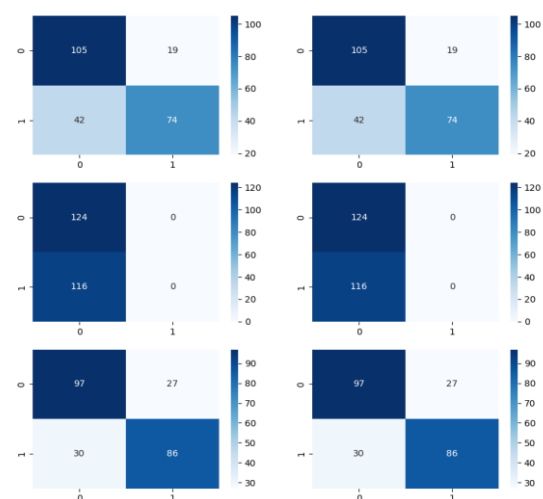
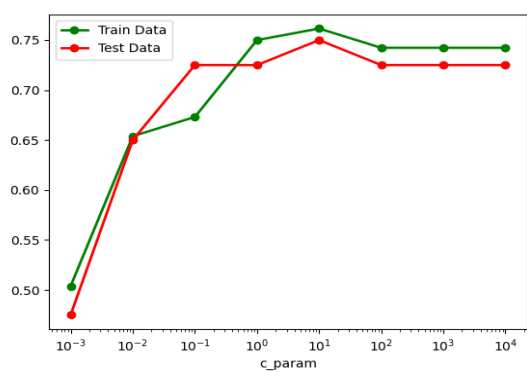
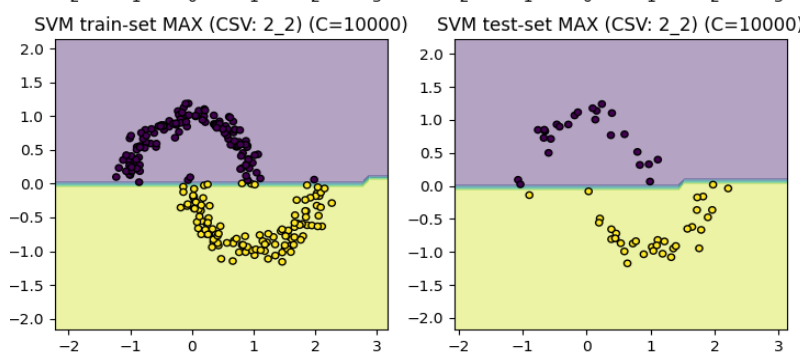
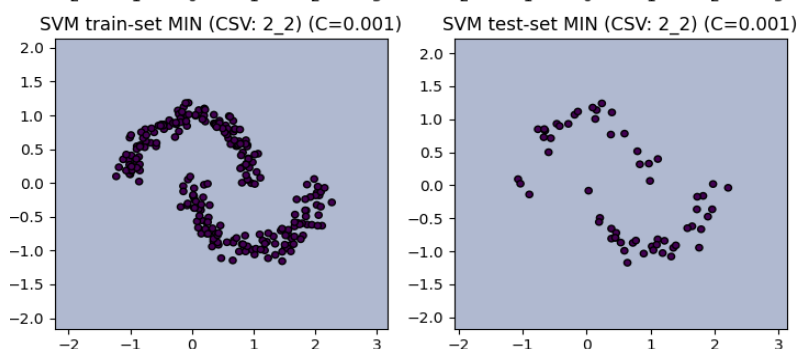
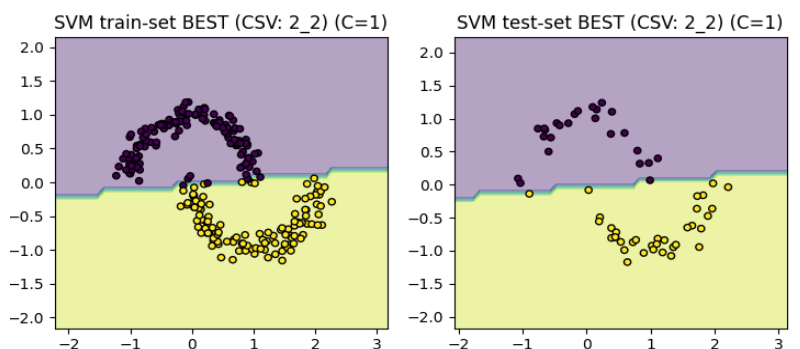
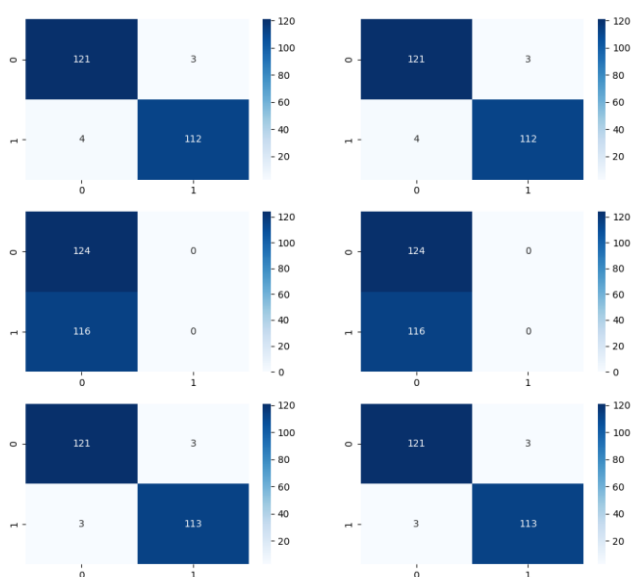
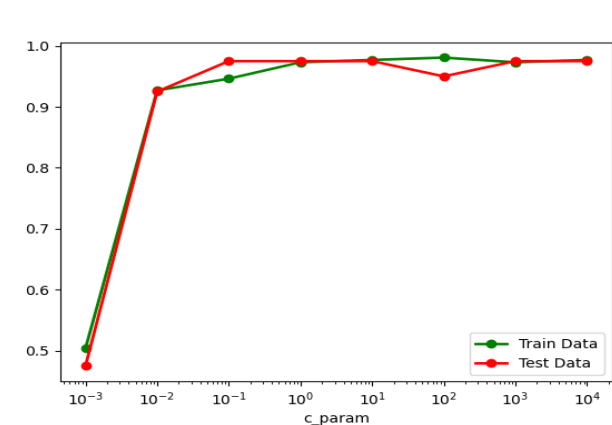
MLP



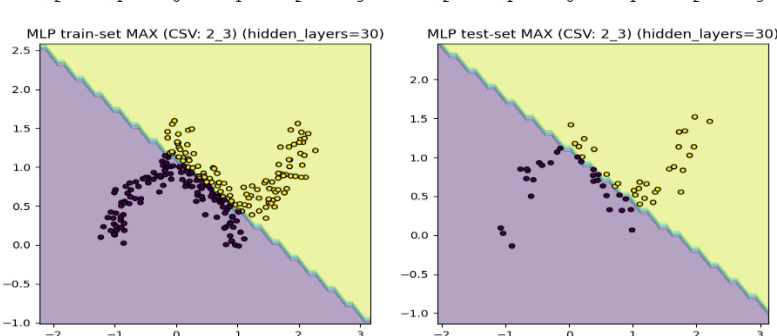
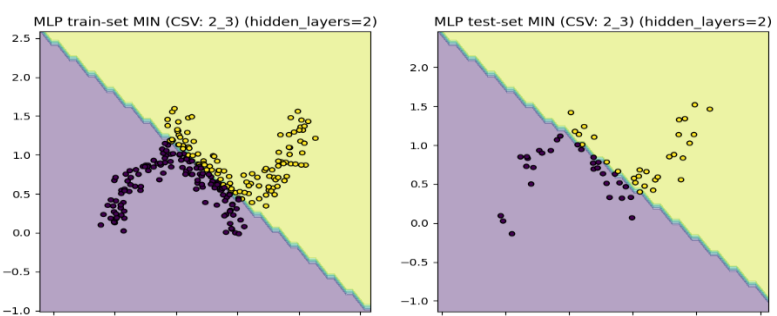
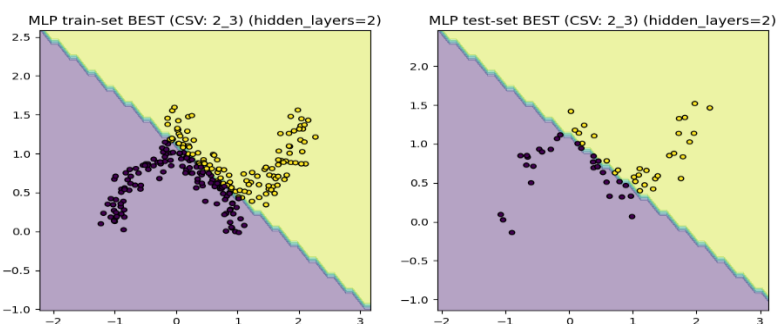
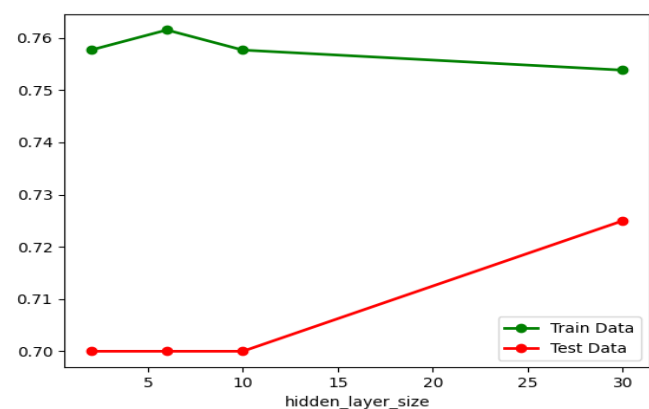
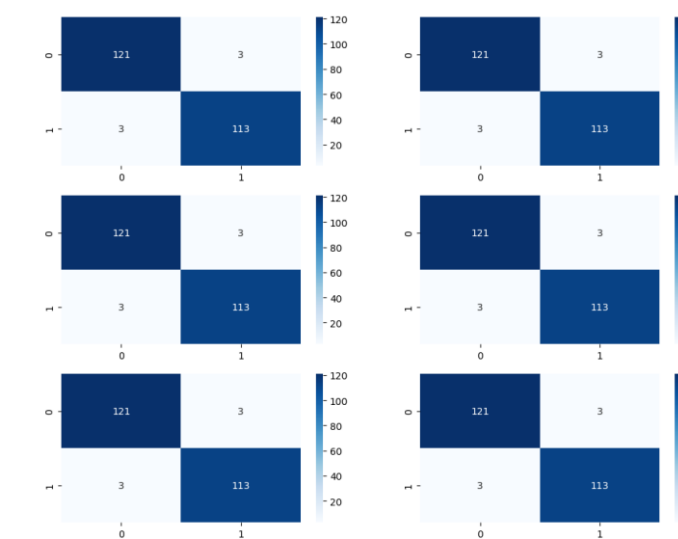
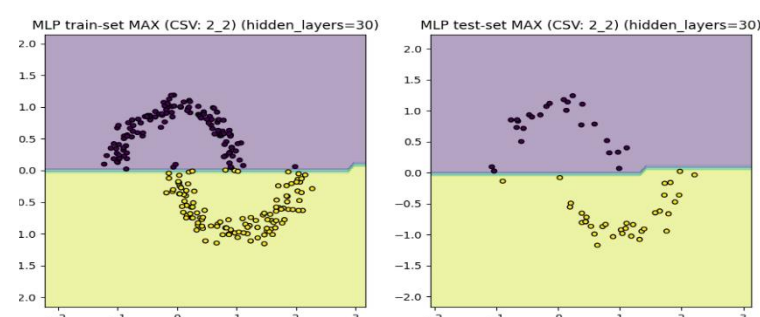
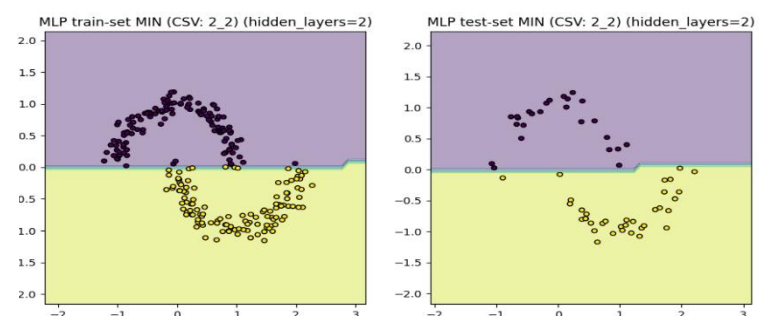
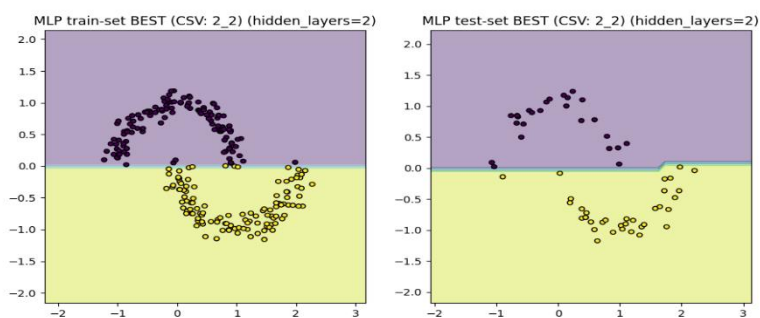
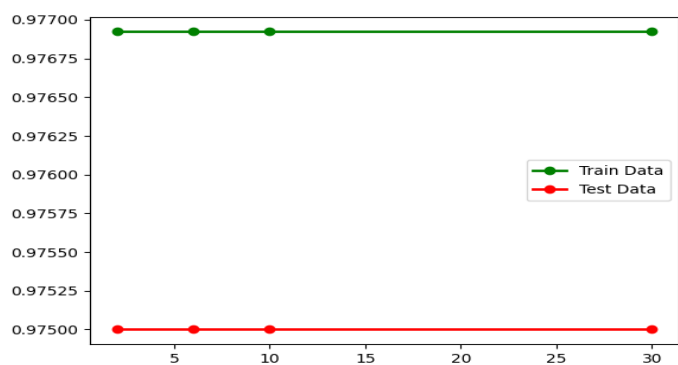
2.2 Wyniki drugiego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz metody K-NN.



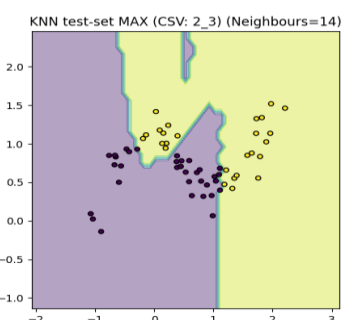
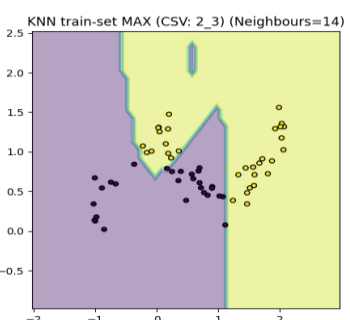
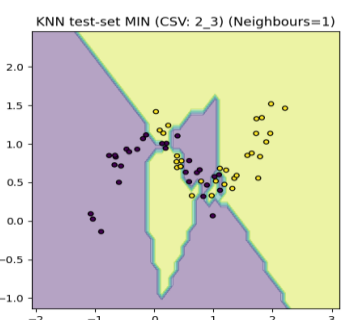
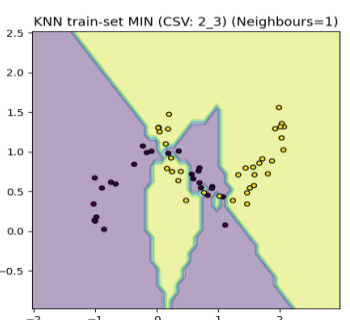
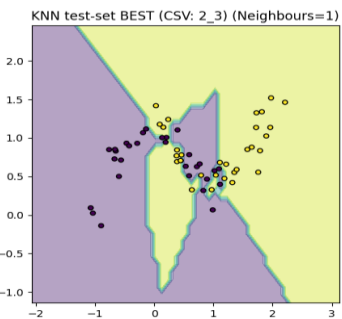
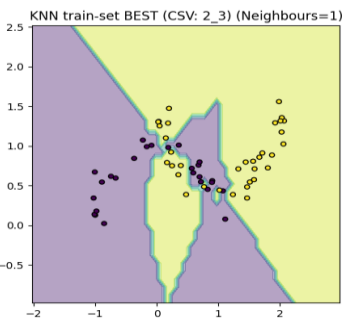
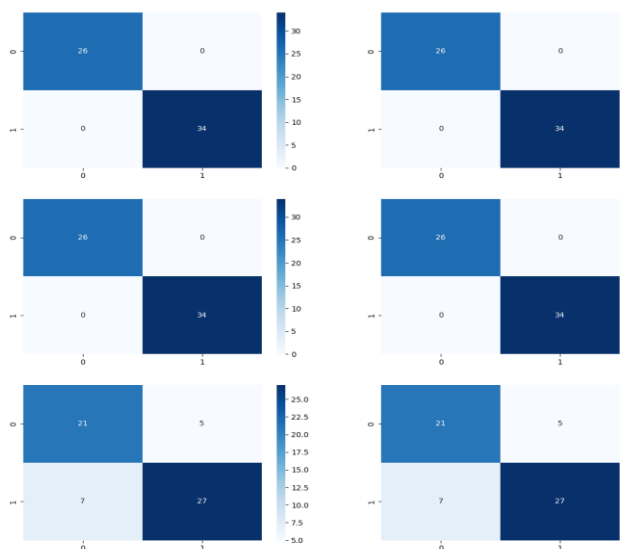
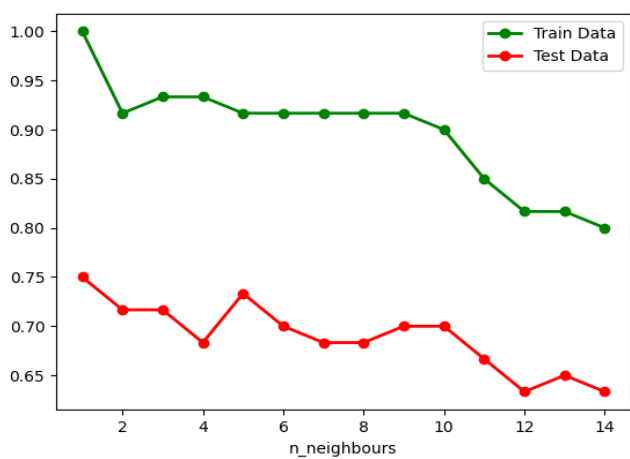
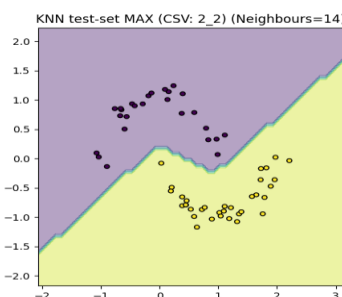
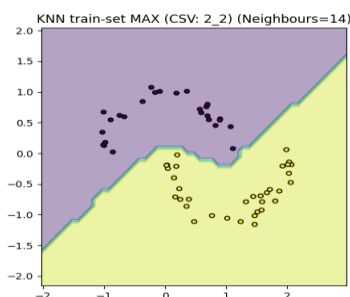
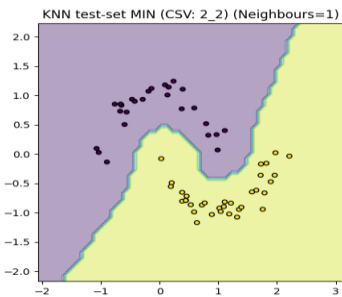
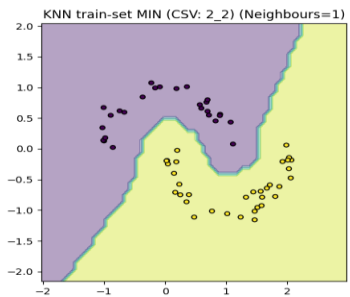
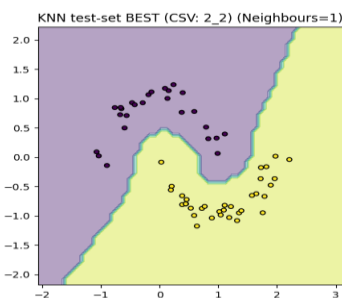
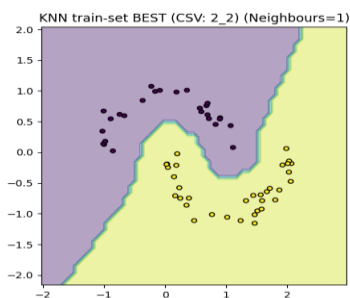
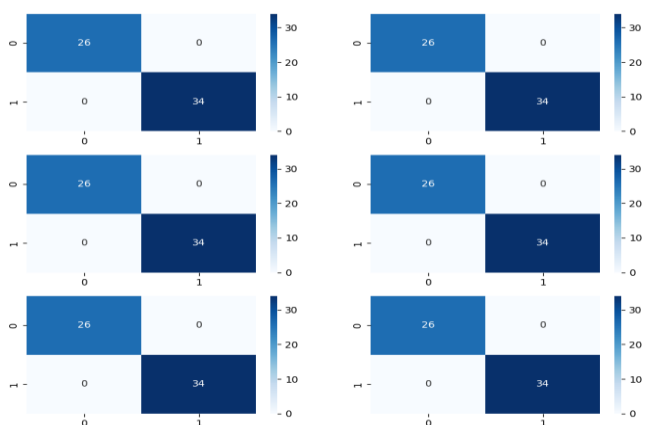
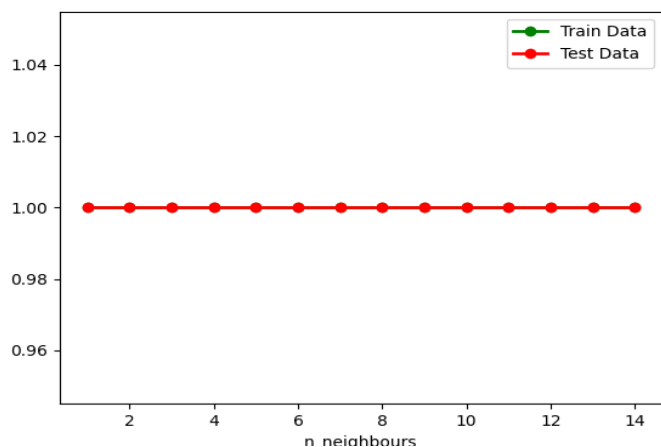
2.3 Wyniki drugiego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz metody SVM.



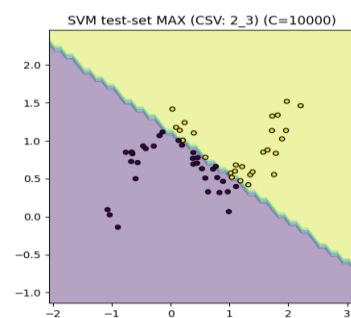
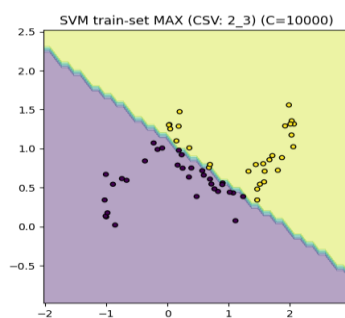
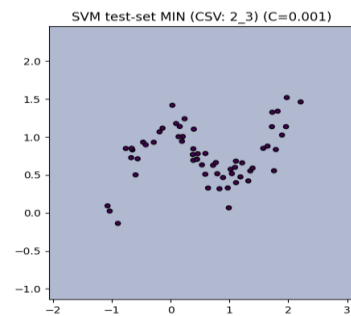
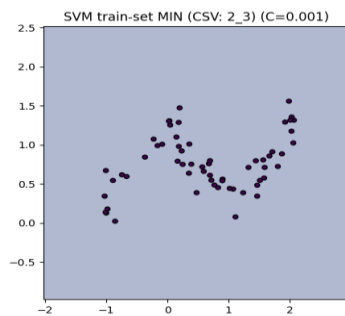
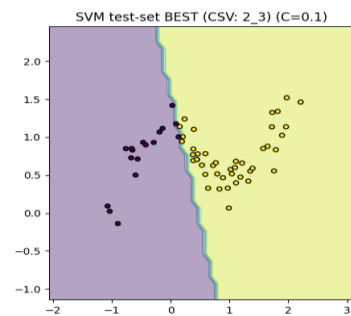
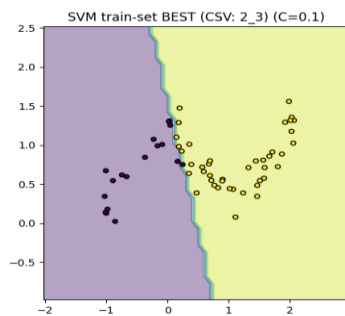
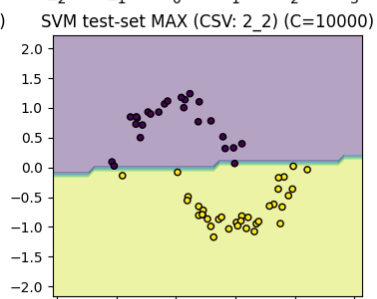
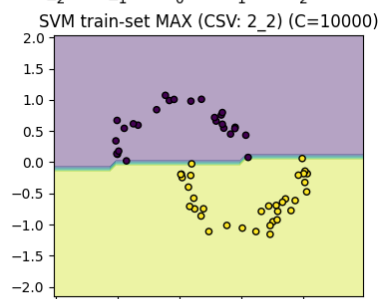
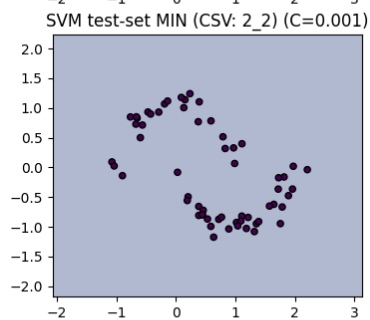
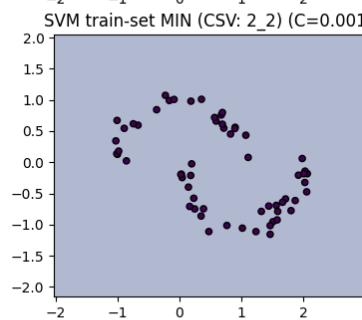
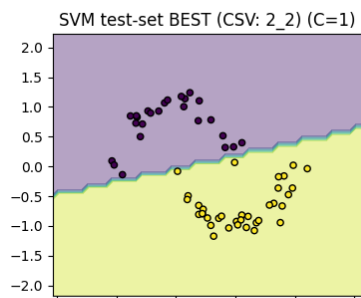
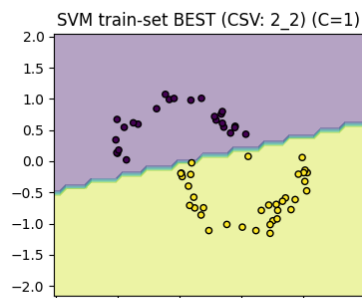
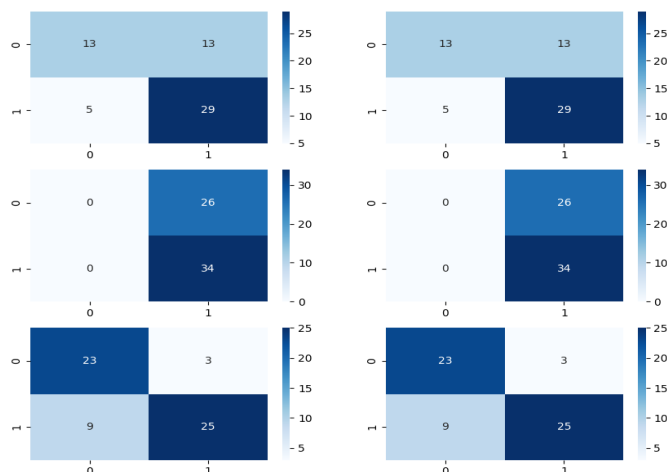
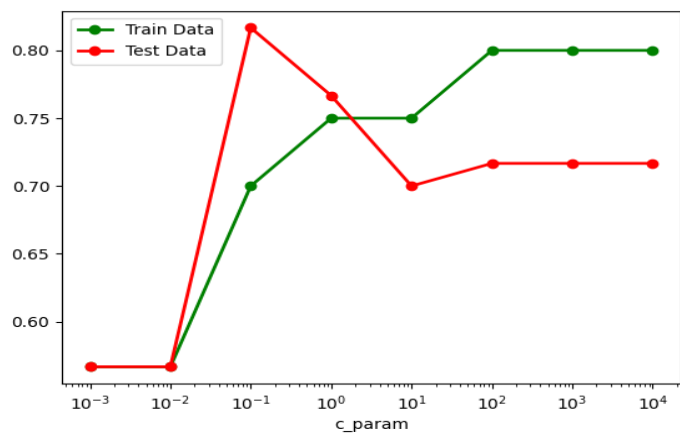
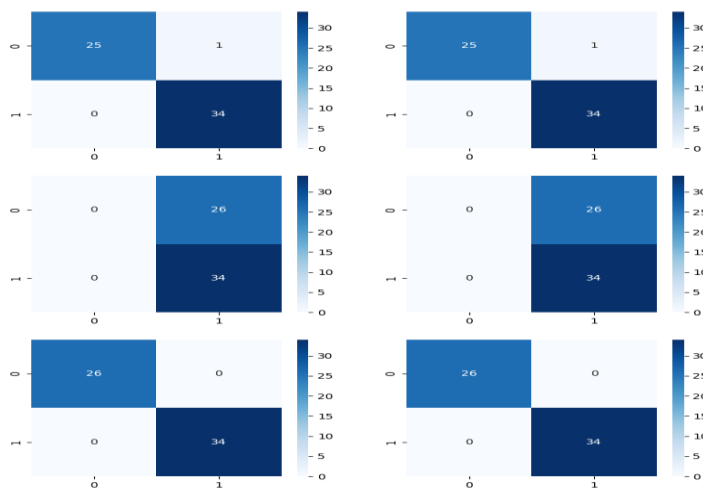
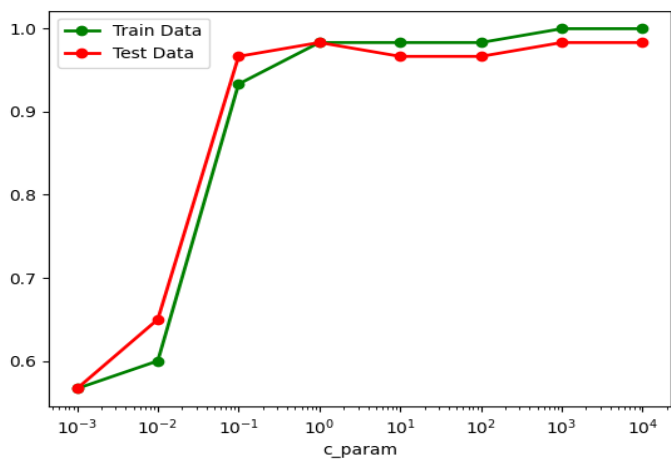
2.4 Wyniki drugiego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz sieci MLP.



2.5 Wyniki trzeciego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz metody K-NN.

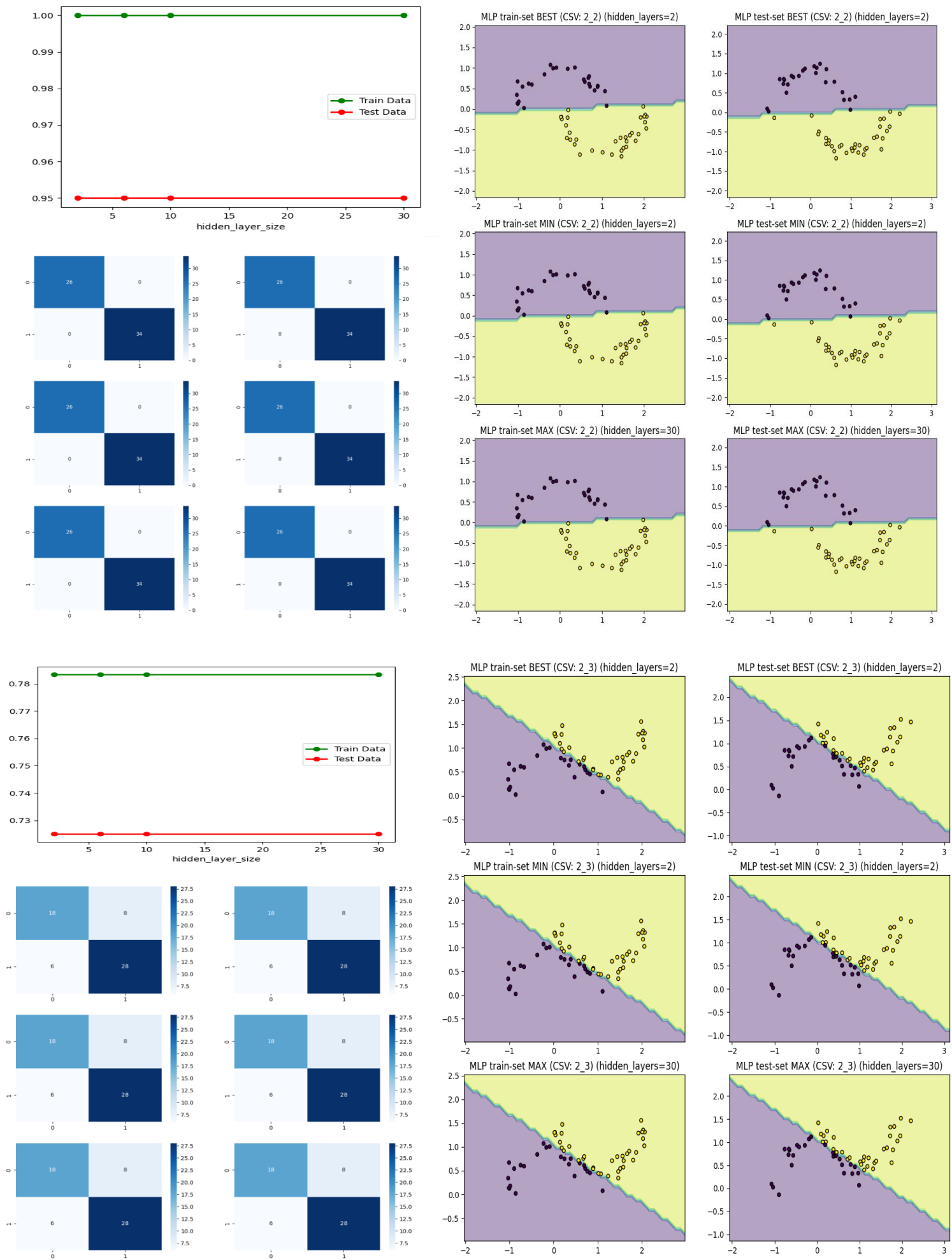


2.6 Wyniki trzeciego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz metody SVM.

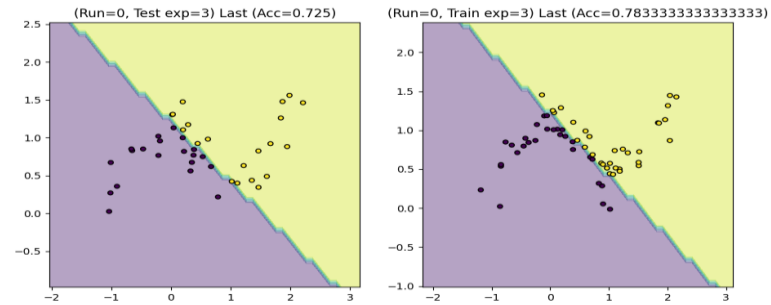
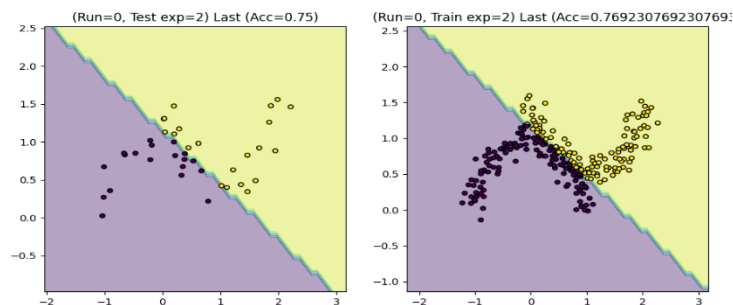
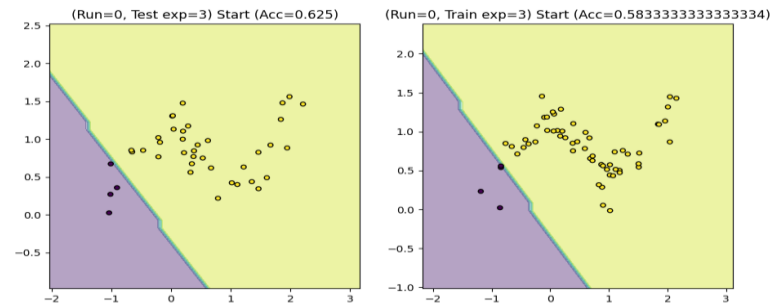
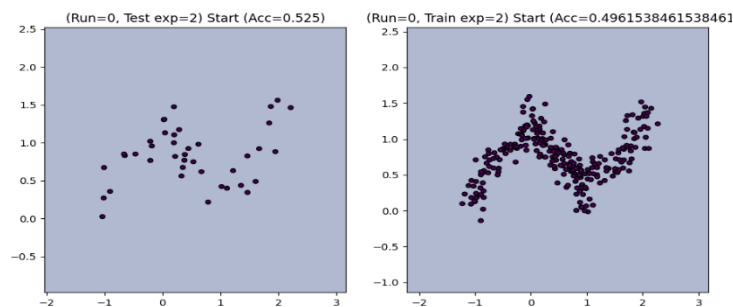
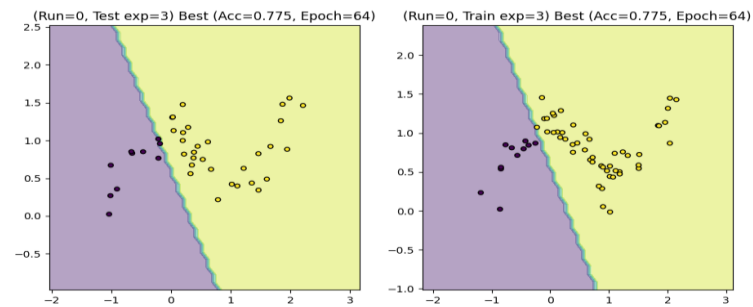
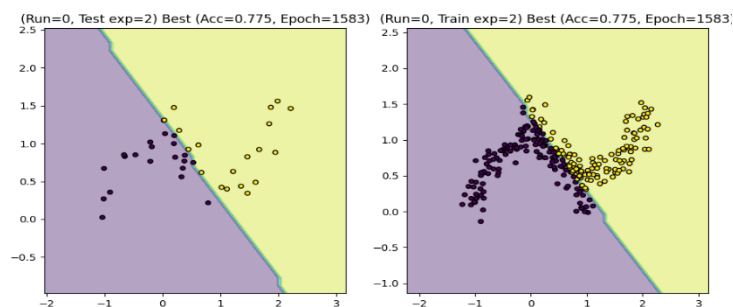
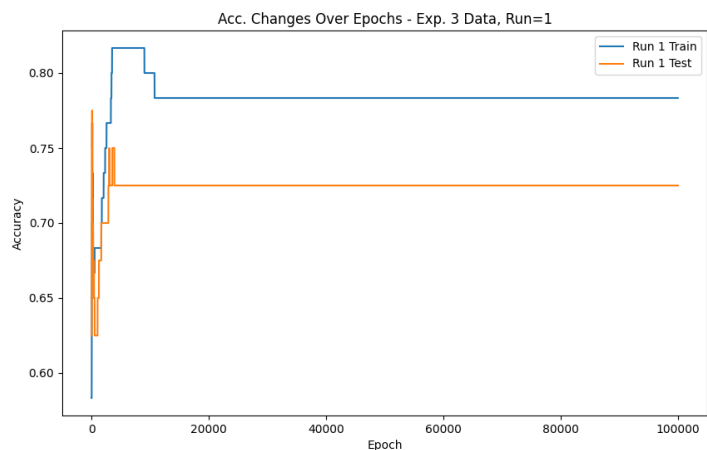
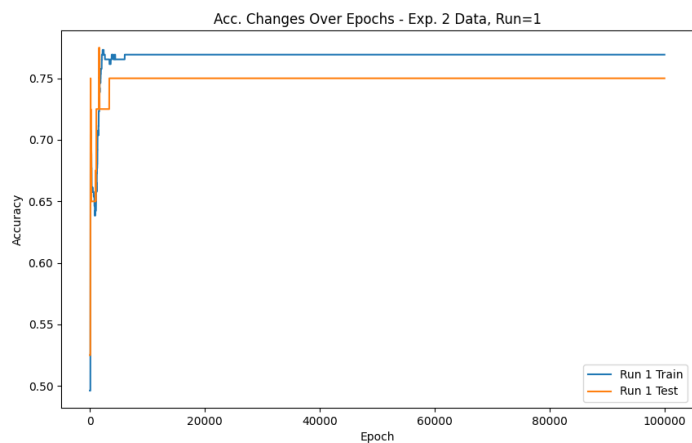


2.7 Wyniki trzeciego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz sieci MLP.

A



2.8 Wyniki czwartego eksperymentu dla sztucznie wygenerowanego zbioru danych 2_3 oraz sieci MLP.



TEST				
Run	AccStart	AccBest	BestEpoch	AccEnd
1.0	0.525	0.775	1583.0	0.75
2.0	0.525	0.775	3108.0	0.75
3.0	0.475	0.75	35.0	0.725
4.0	0.475	0.8	114.0	0.725
5.0	0.6	0.775	41.0	0.725
6.0	0.475	0.775	54.0	0.75
7.0	0.6	0.825	32.0	0.7
8.0	0.45	0.75	91.0	0.75
9.0	0.3	0.775	51.0	0.75
10.0	0.725	0.775	95.0	0.75

TEST				
Run	AccStart	AccBest	BestEpoch	AccEnd
1.0	0.625	0.775	64.0	0.725
2.0	0.225	0.775	132.0	0.725
3.0	0.525	0.75	3583.0	0.725
4.0	0.475	0.75	62.0	0.725
5.0	0.325	0.775	132.0	0.725
6.0	0.375	0.775	70.0	0.725
7.0	0.525	0.775	117.0	0.725
8.0	0.775	0.775	0.0	0.725
9.0	0.475	0.775	111.0	0.725
10.0	0.525	0.775	138.0	0.725

TRAIN				
Run	AccStart	AccBest	BestEpoch	AccEnd
1.0	0.525	0.775	1583.0	0.75
2.0	0.525	0.775	3108.0	0.75
3.0	0.475	0.75	35.0	0.725
4.0	0.475	0.8	114.0	0.725
5.0	0.6	0.775	41.0	0.725
6.0	0.475	0.775	54.0	0.75
7.0	0.6	0.825	32.0	0.7
8.0	0.45	0.75	91.0	0.75
9.0	0.3	0.775	51.0	0.75
10.0	0.725	0.775	95.0	0.75
10.0	0.738	0.775	95.0	0.75

TRAIN				
Run	AccStart	AccBest	BestEpoch	AccEnd
1.0	0.625	0.775	64.0	0.725
2.0	0.225	0.775	132.0	0.725
3.0	0.525	0.75	3583.0	0.725
4.0	0.475	0.75	62.0	0.725
5.0	0.325	0.775	132.0	0.725
6.0	0.375	0.775	70.0	0.725
7.0	0.525	0.775	117.0	0.725
8.0	0.775	0.775	0.0	0.725
9.0	0.475	0.775	111.0	0.725
10.0	0.525	0.775	138.0	0.725
10.0	0.516	0.775	138.0	0.783

2.9 Opis wniosków z eksperymentów przeprowadzonych na sztucznie wygenerowanych zbiorach.

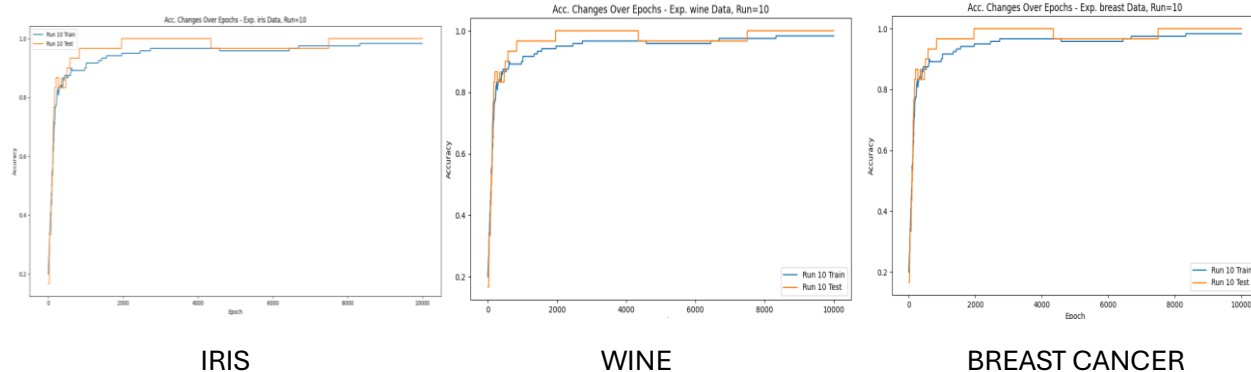
Wynikową granicą decyzyjną klasyfikatora SVC jest hiperpłaszczyzna, więc poprzez konstrukcję przypisać danym może jedynie dwa klastry bez dodatkowych modyfikacji uogólniających. Ze względu jednak na łatwość geometryczną problemu skonstruowania płaszczyzny teoretyczny dolny limit punktów danych koniecznych do znalezienia hiperpłaszczyzny dającej wysoką trafność jest bardzo niski, jednak dane te muszą być dobrane w reprezentatywny sposób co i tak wymaga przeanalizowania zbioru. Pominąć to można podaniem wystarczającej ilości danych by osiągnąć reprezentatywność. Po przekroczeniu tej granicy zwiększanie ilości danych treningowych będzie miało nieznaczny wpływ o ile dane te zachowują rozkład zgodny z rozkładem całego zbioru. Parametr C , będący odwrotnie proporcjonalny do rozmiarów obszaru niepewności klasyfikatora, przy dobrze odseparowanych zbiorach danych nie ma znaczenia dopóki pozostanie powyżej pewnej niskiej wartości krytycznej, przy której obszar niepewności zaczyna zawierać znaczną część podzbiorów klas. W przypadku przemieszania danych parametr C nie może być zbyt wysoki, bo zagwarantuje to pomyłki dla danych znajdujących się w problematycznym podzbiorze. Parametr C nie może jednak być zbyt niski, skutki tego zaobserwować można w wynikach eksperymentów dla parametru $C=0.001$ co skutkuje utworzeniem obszaru niepewności pokrywającego cały zbiór danych i klasyfikowaniem wszystkiego do jednego klastra.

Klasyfikator *k nearest neighbours* przechowuje kopię zbioru treningowego i przypisuje danym testowym etykiety na podstawie większości etykiet danych treningowych w otoczeniu punktu testowego. Kształt tego otoczenia zależy przede wszystkim od wybranej metryki, w przypadku metryki-2 jest to hiperkula o środku w punkcie testowym. Rozmiar otoczenia jest dobierany algorytmicznie aby spełnić warunek ilości sąsiadów. Ze względu na działanie teoretycznym maksymalnym limitem ilości sąsiadów jest ilość punktów treningowych. Taki klasyfikator będzie jednak równoznaczny z przypisaniem każdemu punktowi testowemu etykiety najczęściej występującej w zbiorze treningowym, co mija się z paradygmatem uczenia maszynowego. Zbyt mała ilość sąsiadów również nie powinna wpływać na trafność klasyfikacji dobrze odseparowanych zbiorów, jednak w przypadku zbiorów przemieszanych punkt testowy może trafić w lokalne skupisko danych o mniejszościowej etykiecie. Najlepsza wartość tego parametru będzie więc związana z gęstością zbioru treningowego w sektorze dużego przemieszania. Spodziewane zachowanie pogarszania trafności klasyfikacji pokrywa się z wynikami uzyskanymi w eksperymentach. Dla zbioru 2_3 gdzie dane są przemieszane zaobserwować również można wzrost trafności przy przeskoku z 1 sąsiada do 2 przy zbiorze testowym, jednak dla zbioru treningowego najwyższa wartość osiągnięta jest dla 1, co również pokrywa się ze spodziewanym zachowaniem.

Klasyfikator *Multi-Layer Perceptron* jest rodzajem sieci neuronowej, która składa się z warstw wejściowej, jednej lub więcej warstw ukrytych (parametr `hidden_layer`) oraz warstwy wyjściowej. Każda warstwa składa się z neuronów, które połączone są ze sobą za pomocą wag, każdy neuron wykonuje obliczenia na danych wejściowych a następnie przekazuje je do kolejnej warstwy. Do warstwy wejściowej przekazywany jest wektor cech, a każdy neuron w niej odpowiada jednej cenie, a jego wartość jest równa wartości danej cechy. Warstwy ukryte przetwarzają dane wejściowe poprzez wykonywanie operacji na wartościach neuronów w poprzednich warstwach, w każdej z nich neurony obliczają sumę wejść, którą następnie przekazują przez funkcję aktywacji np. funkcję sigmoidalną czy *Rectified Linear Unit (ReLU)* – te operacje wykonywane są iteracyjnie przez kolejne warstwy ukryte, co pozwala na tworzenie złożonych reprezentacji danych. W warstwie wyjściowej generowane są końcowe wyniki klasyfikacji – zazwyczaj zawiera ona tyle neuronów, ile jest możliwych klas w problemie klasyfikacji. Każdy neuron reprezentuje prawdopodobieństwo przynależności do danej klasy. Zwiększanie ilości warstw ukrytych za pomocą parametru `hidden_layer` pozwala na dostosowanie modelu do złożoności problemu, jednakże może zwiększyć ryzyko przeuczenia modelu – zwłaszcza kiedy ilość danych treningowych jest ograniczona, może to przyczynić się do obniżenia dokładności na danych testowych. Dla zbioru 2_3 w eksperymencie nr.2, dla dużej ilości danych treningowych dokładność wzrasta wraz ze zwiększaniem parametru `hidden_layer`, natomiast dla eksperymentu nr.3, dla małej ilości danych treningowych nie zaobserwowano wpływu tego parametru na zmianę dokładności poprawnie przypisanych etykiet.

2.10 Opis działania analizowanych metod klasyfikacji w przypadku rzeczywistych zbiorów danych.

- Trenowanie klasyfikatora MLP dla rzeczywistych danych iris, wine oraz breast cancer,



IRIS TRAIN					WINE TRAIN					BREAST TRAIN				
Run	AccStart	AccBest	BestEpoch	AccEnd	Run	AccStart	AccBest	BestEpoch	AccEnd	Run	AccStart	AccBest	BestEpoch	AccEnd
1,00	0,23	1,00	1137,00	0,97	1,00	0,33	1,00	1074,00	1,00	1,00	0,38	0,98	797,00	0,98
2,00	0,60	1,00	1672,00	1,00	2,00	0,39	1,00	1711,00	1,00	2,00	0,34	0,99	2173,00	0,98
3,00	0,03	1,00	2431,00	1,00	3,00	0,36	1,00	734,00	1,00	3,00	0,38	0,99	344,00	0,97
4,00	0,10	1,00	2620,00	1,00	4,00	0,39	1,00	266,00	1,00	4,00	0,61	0,99	283,00	0,99
5,00	0,27	1,00	4075,00	0,97	5,00	0,08	1,00	1165,00	1,00	5,00	0,37	0,99	147,00	0,98
6,00	0,23	0,97	171,00	0,97	6,00	0,31	1,00	2263,00	1,00	6,00	0,38	0,98	1004,00	0,98
7,00	0,27	1,00	1676,00	1,00	7,00	0,25	1,00	162,00	1,00	7,00	0,30	0,99	749,00	0,97
8,00	0,53	1,00	571,00	1,00	8,00	0,22	1,00	425,00	1,00	8,00	0,62	0,98	182,00	0,98
9,00	0,10	1,00	2165,00	1,00	9,00	0,33	1,00	7815,00	1,00	9,00	0,37	0,98	1945,00	0,97
10,00	0,17	1,00	1959,00	1,00	10,00	0,56	1,00	7455,00	1,00	10,00	0,48	0,99	319,00	0,99
10,00	0,20	1,00	1959,00	0,98	10,00	0,51	1,00	7455,00	1,00	10,00	0,52	0,99	319,00	0,99

IRIS TEST					WINE TEST					BREAST TEST				
Run	AccStart	AccBest	BestEpoch	AccEnd	Run	AccStart	AccBest	BestEpoch	AccEnd	Run	AccStart	AccBest	BestEpoch	AccEnd
1,00	0,23	1,00	1137,00	0,97	1,00	0,33	1,00	1074,00	1,00	1,00	0,38	0,98	797,00	0,98
2,00	0,60	1,00	1672,00	1,00	2,00	0,39	1,00	1711,00	1,00	2,00	0,34	0,99	2173,00	0,98
3,00	0,03	1,00	2431,00	1,00	3,00	0,36	1,00	734,00	1,00	3,00	0,38	0,99	344,00	0,97
4,00	0,10	1,00	2620,00	1,00	4,00	0,39	1,00	266,00	1,00	4,00	0,61	0,99	283,00	0,99
5,00	0,27	1,00	4075,00	0,97	5,00	0,08	1,00	1165,00	1,00	5,00	0,37	0,99	147,00	0,98
6,00	0,23	0,97	171,00	0,97	6,00	0,31	1,00	2263,00	1,00	6,00	0,38	0,98	1004,00	0,98
7,00	0,27	1,00	1676,00	1,00	7,00	0,25	1,00	162,00	1,00	7,00	0,30	0,99	749,00	0,97
8,00	0,53	1,00	571,00	1,00	8,00	0,22	1,00	425,00	1,00	8,00	0,62	0,98	182,00	0,98
9,00	0,10	1,00	2165,00	1,00	9,00	0,33	1,00	7815,00	1,00	9,00	0,37	0,98	1945,00	0,97
10,00	0,17	1,00	1959,00	1,00	10,00	0,56	1,00	7455,00	1,00	10,00	0,48	0,99	319,00	0,99

- Klasyfikacja KNeighborsClassifier dla rzeczywistych zbiorów danych – macierz pomyłek oraz przebieg wartości accuracy w zależności od wartości parametru n_neighbors

