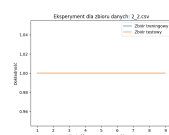
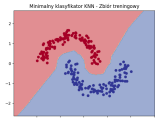


Obliczenia Inteligentne		Projekt 1
Grupa: Grupa 1	Dzień i czas: Czwartek, 10:00	Rok akademicki: 2023/24
Imię i nazwisko: JAKUB PAWLAK		Imię i nazwisko: MAGDALENA PAKUŁA

1 strona — Wyniki pierwszego eksperymentu dla trzech sztucznie wygenerowanych zbiorów danych 2_1, 2_2 oraz 2_3. Dla każdego zbioru i metody (klasyfikator SVM z różnymi wartościami parametru kernel oraz sieć MLP z różnymi wartościami parametru activation) należy przygotować wizualizacje przebiegu granicy decyzyjnej dla przypadku, gdy accuracy na zbiorze treningowym ma najlepszą wartość (zmianie podlegają parametr c dla klasyfikatora SVM i liczba neuronów w warstwie ukrytej dla sieci MLP).



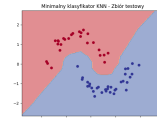
(a) Dokładność



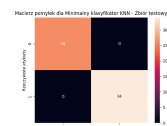
(b) Min, train



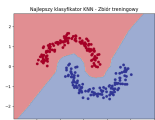
(c) Min, train



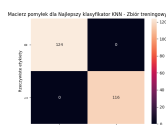
(d) Min, test



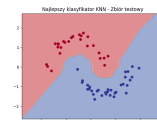
(e) Min, test



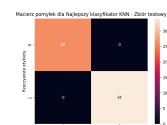
(f) Best, train



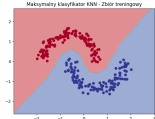
(g) Best, train



(h) Best, test



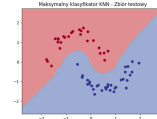
(i) Best, test



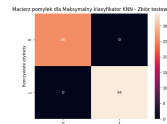
(j) Max, train



(k) Max, train

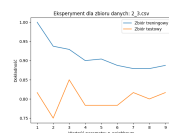


(l) Max, test

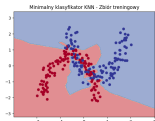


(m) Max, test

Rysunek 1: Wyniki dla zbioru 2_2



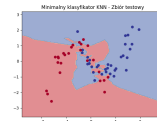
(a) Dokładność



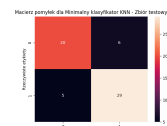
(b) Min, train



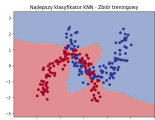
(c) Min, train



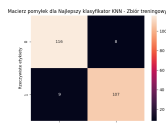
(d) Min, test



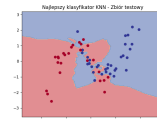
(e) Min, test



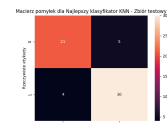
(f) Best, train



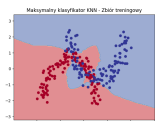
(g) Best, train



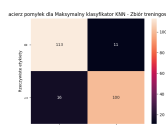
(h) Best, test



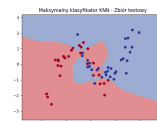
(i) Best, test



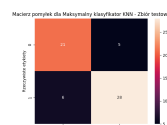
(j) Max, train



(k) Max, train



(l) Max, test



(m) Max, test

Rysunek 2: Wyniki dla zbioru 2_3

3 strona — Wyniki drugiego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz metody SVM. Dla każdego zbioru należy pokazać wykres obrazujący zmianę wartości accuracy na zbiorach treningowym i testowym przy zmieniającym się parametrze C oraz wizualizację przebiegu granicy decyzyjnej na zbiorach treningowym i testowym dla: najmniejszej, najlepszej (wartość accuracy na zbiorze testowym) i największej wartości tego parametru. Dodatkowo przy każdej wizualizacji należy pokazać jak wygląda macierz pomyłek. Wartości parametru C powinny się zmieniać wykładniczo, a na wykresie dobrze jest zastosować skalę logarytmiczną.

4 strona — Wyniki drugiego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz sieci MLP. Dla każdego zbioru należy pokazać wykres obrazujący zmianę wartości accuracy na zbiorach treningowym i testowym przy zmieniającej się liczbie neuronów w warstwie ukrytej oraz wizualizację przebiegu granicy decyzyjnej na zbiorach treningowym i testowym dla: najmniejszej, najlepszej (wartość accuracy na zbiorze testowym) i największej wartości tego parametru. Dodatkowo przy każdej wizualizacji należy pokazać jak wygląda macierz pomyłek.

5 strona — Wyniki trzeciego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz metody K-NN. Dla każdego zbioru należy pokazać wykres obrazujący zmianę wartości accuracy na zbiorach treningowym i testowym przy zmieniającym się parametrze `n_neighbours` oraz wizualizację przebiegu granicy decyzyjnej na zbiorach treningowym i testowym dla: najmniejszej, najlepszej (wartość accuracy na zbiorze testowym) i największej wartości tego parametru. Dodatkowo przy każdej wizualizacji należy pokazać jak wygląda macierz pomyłek.

6 strona — Wyniki trzeciego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz metody SVM. Dla każdego zbioru należy pokazać wykres obrazujący zmianę wartości accuracy na zbiorach treningowym i testowym przy zmieniającym się parametrze C oraz wizualizację przebiegu granicy decyzyjnej na zbiorach treningowym i testowym dla: najmniejszej, najlepszej (wartość accuracy na zbiorze testowym) i największej wartości tego parametru. Dodatkowo przy każdej wizualizacji należy pokazać jak wygląda macierz pomyłek. Wartości parametru C powinny się zmieniać wykładniczo, a na wykresie dobrze jest zastosować skalę logarytmiczną.

7 strona — Wyniki trzeciego eksperymentu dla dwóch sztucznie wygenerowanych zbiorów danych 2_2 i 2_3 oraz sieci MLP. Dla każdego zbioru należy pokazać wykres obrazujący zmianę wartości accuracy na zbiorach treningowym i testowym przy zmieniającej się liczbie neuronów w warstwie ukrytej oraz wizualizację przebiegu granicy decyzyjnej na zbiorach treningowym i testowym dla: najmniejszej, najlepszej (wartość accuracy na zbiorze testowym) i największej wartości tego parametru. Dodatkowo przy każdej wizualizacji należy pokazać jak wygląda macierz pomyłek.

8 strona — Wyniki czwartego eksperymentu dla sztucznie wygenerowanego zbioru danych 2_3 oraz sieci MLP. Dla rozważanego zbioru należy rozważyć przypadki z różną liczbą danych treningowych (parametr `train_size` równy różnym wartościom użytym odpowiednio w eksperymentach drugim i trzecim). Dla obu przypadków należy zaprezentować wykres zmian `accuracy` na zbiorach treningowym i testowym w kolejnych epokach oraz wizualizację przebiegu granicy decyzyjnej na zbiorach treningowym i testowym dla epoki: zerowej (przed rozpoczęciem nauki), najlepszej (wartość `accuracy` na zbiorze testowym) i ostatniej (po zakończeniu nauki). Dodatkowo w każdym z przypadków należy uruchomić proces treningu 10 razy z różnymi wagami początkowymi i w tabeli zamieścić wartości `accuracy` na zbiorze testowym i treningowym dla epoki: pierwszej (początek nauki), najlepszej (wartość `accuracy` na zbiorze testowym) i ostatniej (po zakończeniu nauki). W przypadku wartości najlepszej należy również podać numer epoki kiedy ją osiągnięto. Liczbę neuronów w warstwie ukrytej należy dobrać jako tą optymalną wynikającą odpowiednio z eksperymentów drugiego i trzeciego.

9 strona — Opis wniosków z eksperymentów przeprowadzonych na sztucznie wygenerowanych zbiorach. W przypadku wszystkich eksperymentów należy zwrócić uwagę na kształt uzyskiwanych granic decyzyjnych i związane z nim zdolności uogólniające poszczególnych rodzajów klasyfikatorów (wpływ hiperparametrów) oraz wpływ liczby danych treningowych. W eksperymencie czwartym należy dodatkowo skupić się na zdolnościach uogólniających w kolejnych epokach nauki oraz na wpływie sposobu zainicjalizowania sieci. Wnioski powinny mieć charakter ogólny, pozwalający przenieść je na przypadek, w którym nie ma możliwości zwizualizowania danych. Każdy wniosek powinien być poparty odniesieniami do wyników przedstawionych na pierwszych czterech stronach raportu.

10 strona — Opis działania analizowanych metod klasyfikacji w przypadku rzeczywistych zbiorów danych. Podczas tworzenia klasyfikatorów warto skorzystać z wniosków wyciągniętych podczas wcześniejszych eksperymentów. Uzyskane wyniki należy zaprezentować w zwartej formie (warto wykorzystać tabele i/lub wykresy), a wnioski należy poprzeć odwołaniami do tych wyników.