

Imię i nazwisko: _____

Rozwiązania zadań powienny zostać spakowany łącznie z zestawem zadań, bazą danych oraz dodatkowymi plikami jako *grupa_imie_nazwisko*. Następnie spakowane pliki należy wysłać na platformę zdalnej edukacji.

Uwaga – kod powinien być skomentowany aby w łatwy sposób można było wiedzieć gdzie i co jest.

1. Pobierz bazę danych o nazwie *Wine Quality Dataset*. Następnie zaimplementuj klasę do obsługi danych: metoda *ShuffleData()* służąca do tasowania zbioru, *SplitData()* przeznaczona do dzielenia zbioru na zbiór treningowy/walidacyjny w proporcji $x : y$, gdzie x oraz y są argumentami metody ($x + y = 100$) [5 punktów].

Zaimplementuj metodę *NormalizeData()*, wykonującą normalizację atrybutów korzystając ze standaryzacji:

$$x = \frac{x - \mu}{\sigma}, \quad (1)$$

gdzie μ to średnia danego atrybutu, a σ to odchylenie standardowe danego atrybutu [5 punktów].

Powyższa klasa powinna być przetestowana.

2. Zaimplementuj algorytm k najbliższych sąsiadów dla bazy danych *Wine Quality Dataset* przy użyciu metryki Minkowskiego z $m = 2$. Następnie wykonaj testy pomiarów dokładności metody dla danych znormalizowanych/nieznormalizowanych przy zastosowaniu $k = 4$. Otrzymane wyniki powinny być udokumentowane zrzutem ekranu [10 punktów].
3. Zaproponuj i zaimplementuj budowę klasyfikatora miękkiego dla bazy *Wine Quality Dataset*. Dokonaj pomiarów dokładności na zbiorze walidacyjnym. [10 punktów].
4. Zaimplementuj system Takagi-Sugeno z użyciem trapezoidowych funkcji przynależności i metody FOM (first-of-max) jako wyostrzenia. Sterownik rozmyty powinien oceniać jakość życia w zadanym mieście biorąc pod uwagę następujące informacje pozyskane przez czujniki temperatury, wilgotności, ciśnienia atmosferycznego i zanieczyszczenia powietrza. Uzasadnij dobór parametrów w funkcjach, oraz przetestuj dla kilku różnych przypadków. [10 punktów].