

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 *Wheat Seeds 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 min-max [5 punktów].

Powyższa klasa powinna być przetestowana.

2. Zaimplementuj klasyfikator Bayesa dla bazy danych *Wheat Seeds Dataset*. Następnie wykonaj testy pomiarów dokładności metody dla danych znormalizowanych/nieznormalizowanych. Otrzymane wyniki powinny być udokumentowane zrzutem ekranu [10 punktów].
3. Zaproponuj i zaimplementuj budowę klasyfikatora miękkiego dla bazy *Wheat Seeds Dataset*. Dokonaj pomiarów dokładności na zbiorze walidacyjnym. [10 punktów].
4. Zaimplementuj system Takagi-Sugeno z użyciem Gaussowskiej 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].