

### **PB14 - Wykorzystanie algorytmu ewolucyjnego do doboru parametrów SVM do zadania klasyfikacji**

Przedmiotem projektu jest wykorzystanie algorytmu ewolucyjnego do optymalizacji parametrów maszyny wektorów nośnych (ang. *support vector machine*, SVM).

SVM to algorytm uczenia maszynowego, tworzący specyficzny model sieci neuronowej, w której pod uwagę brane są jedynie najtrudniej separowalne punkty przestrzeni, zwane wektorami nośnymi. Jej uczenie polega na dobieraniu wag w celu maksymalizacji marginesu, separującego skrajne punkty, definiujące różne klasy. Sieci te dedykowane są głównie zagadnieniom klasyfikacji, w których w sposób jednoznaczny rozdzielają jedną klasę od pozostałych, zwracając wynik z zakresu  $\{-1, 1\}$ . Najprostszym przypadkiem są dane liniowo separowalne, które jednak rzadko występują w problemach rzeczywistych. W celu radzenia sobie z nielinowością granicy decyzyjnej chętnie wykorzystywane jest jądro RBF (ang. *radial base function*). Na jakość rozwiązania duży wpływ ma szerokość  $\sigma$  zastosowanej funkcji bazowej oraz wartość współczynnika regularyzacji  $\alpha$ .

Podejście ewolucyjne to technika optymalizacji globalnej inspirowana biologicznym procesem ewolucji, coraz częściej wykorzystywana w procesie automatycznego uczenia maszynowego. Algorytm ewolucyjny składa się z kilku kroków. Inicjowana jest populacja bazowa  $P(t)$ . Dla każdego z losowo wygenerowanych osobników jest obliczana wartość funkcji przystosowania, czyli miara określająca, jak dobrze dany osobnik jest dopasowany do środowiska. Kolejnym krokiem jest reprodukcja - do  $P(t + 1)$  kopiowane są losowo wybrane osobniki z populacji bazowej (prawdopodobieństwo wylosowania jest wyższe dla osobników o większej wartości funkcji przystosowania). Osobniki z  $P(t + 1)$  poddawane są operacjom genetycznym (najpopularniejszymi są krzyżowanie i mutacja). Następuje sukcesja do następnej generacji. Algorytm jest wykonywany do spełnienia warunku zatrzymania.

Projekt zostanie napisany w języku Python w wersji 3 z użyciem gotowego algorytmu SVM z biblioteki *scikit-learn*. Algorytm ewolucyjny zostanie zaimplementowany samodzielnie. Do testów algorytmu użyte będą zbiory danych dostarczone przez prowadzącego, które zostaną podzielone na zbiory uczące i testowe. Algorytm ewolucyjny będzie optymalizował wspomniane wcześniej parametry  $\sigma$  oraz  $\alpha$  algorytmu SVM. Wyniki przeprowadzonej optymalizacji zostaną ocenione w oparciu o skuteczność klasyfikacji SVM na zbiorze testowym.