

Dokumentacja do WSI - ćwiczenie 3

Regresja i klasyfikacja

Wykonano przez Marfenko Mykhailo

Struktura programu

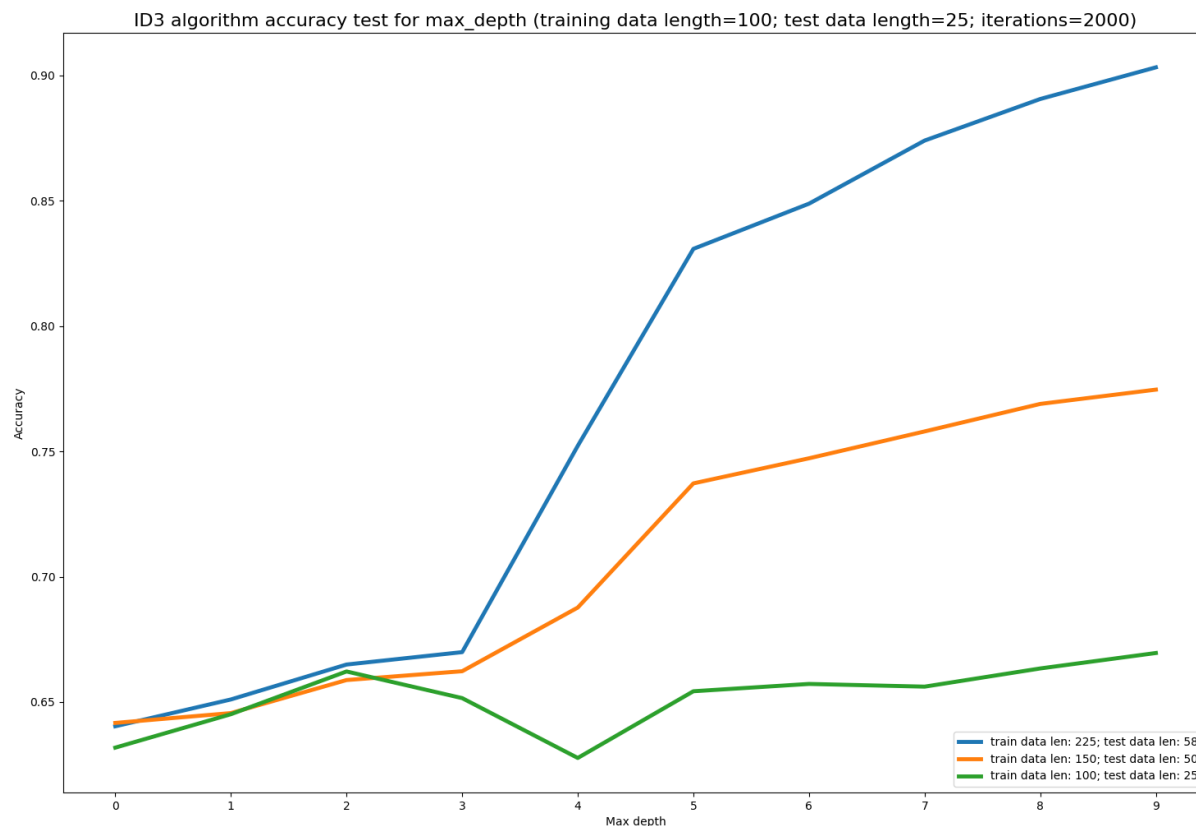
Program jest podzielony na 2 pliki:

Id3.py – Ten plik zawiera funkcjonalność algorytmu ID3 z parametrem `max_depth`. Tutaj znajduje się funkcja `build_tree()` która zwraca wynikowe drzewo decyzyjne jako wielowymiarowy słownik.

experiments.py – skrypt, który bada precyzję (accuracy) drzewa w zależności od parametru `max_depth` oraz długości danych testujących i długości danych trenujących.

Eksperymenty

Opis: Ten eksperyment pokazuje zależność parametru `max_depth` dla drzewa decyzyjnego z różnymi zestawami danych testujących i trenujących.



Wyniki: Nie trudno zauważyć, że im większa głębokość drzewa, tym jest lepsza precyzja, ale to nie zawsze tak jest. Jeżeli podać mały zestaw danych, to istnieje możliwość, że drzewo nie ma wystarczająco danych dla dobrej precyzji nawet z wielką głębokością. Zaś jeżeli danych dla trenowania drzewa jest wystarczająco, to precyzja polepsza wraz z głębokością. Nie jest to zawsze liniowa zależność. To zależy od danych trenujących.