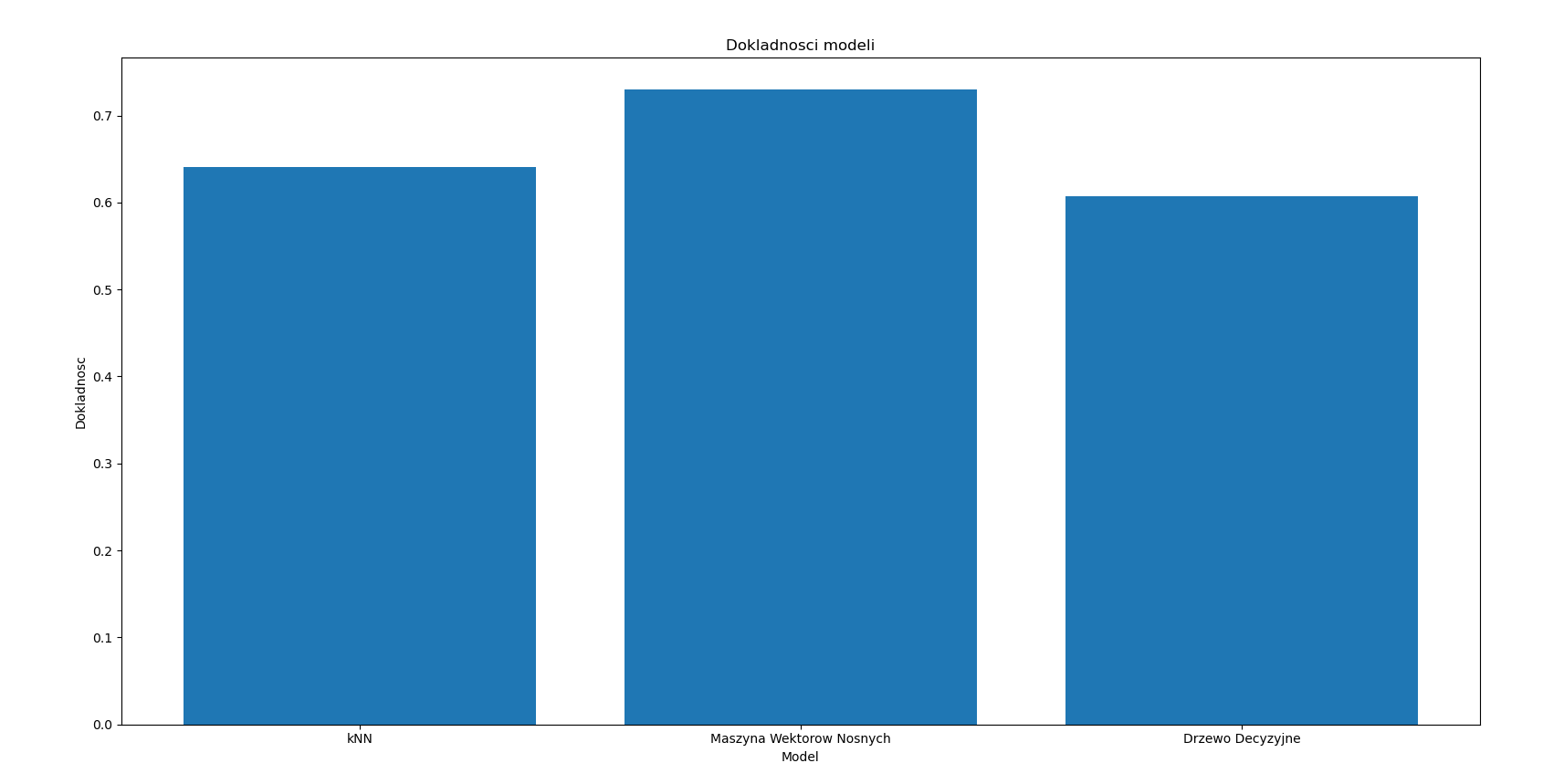
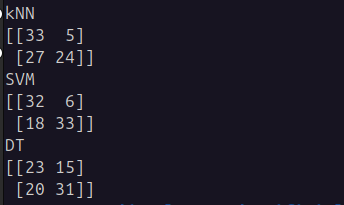
Praca Domowa nr 1

Michał Gagoś

Nasze zastosowanie jest przykładem klasyfikacji z dwoma klasami, gdzie wiele parametrów takich jak wiek, waga itd świadczy o tym czy osoba jest chora na cukrzycę czy nie.

Po przeprowadzeniu badań odnośnie dokładności klasyfikacji wykonywanej przez różne modele sztucznej inteligencji otrzymujemy następujące wyniki.

 1. Porównanie dokładności

 2. Macierz pomylek

Jak widzimy najwyższą dokładnością (ponad 70%) wykazał się model **SVM**.

Ma to dużo sensu gdy weźmiemy pod uwagę jego działanie i nasze dane. Działa on przez dzielenie przestrzeni wielowymiarowej na dwie części. Podział na większą ilość klas jest możliwy, ale wymaga kolejnych iteracji.

Model ten jest też bardzo dobry w generalizowaniu. Przeuczenie go jest mniej prawdopodobne.

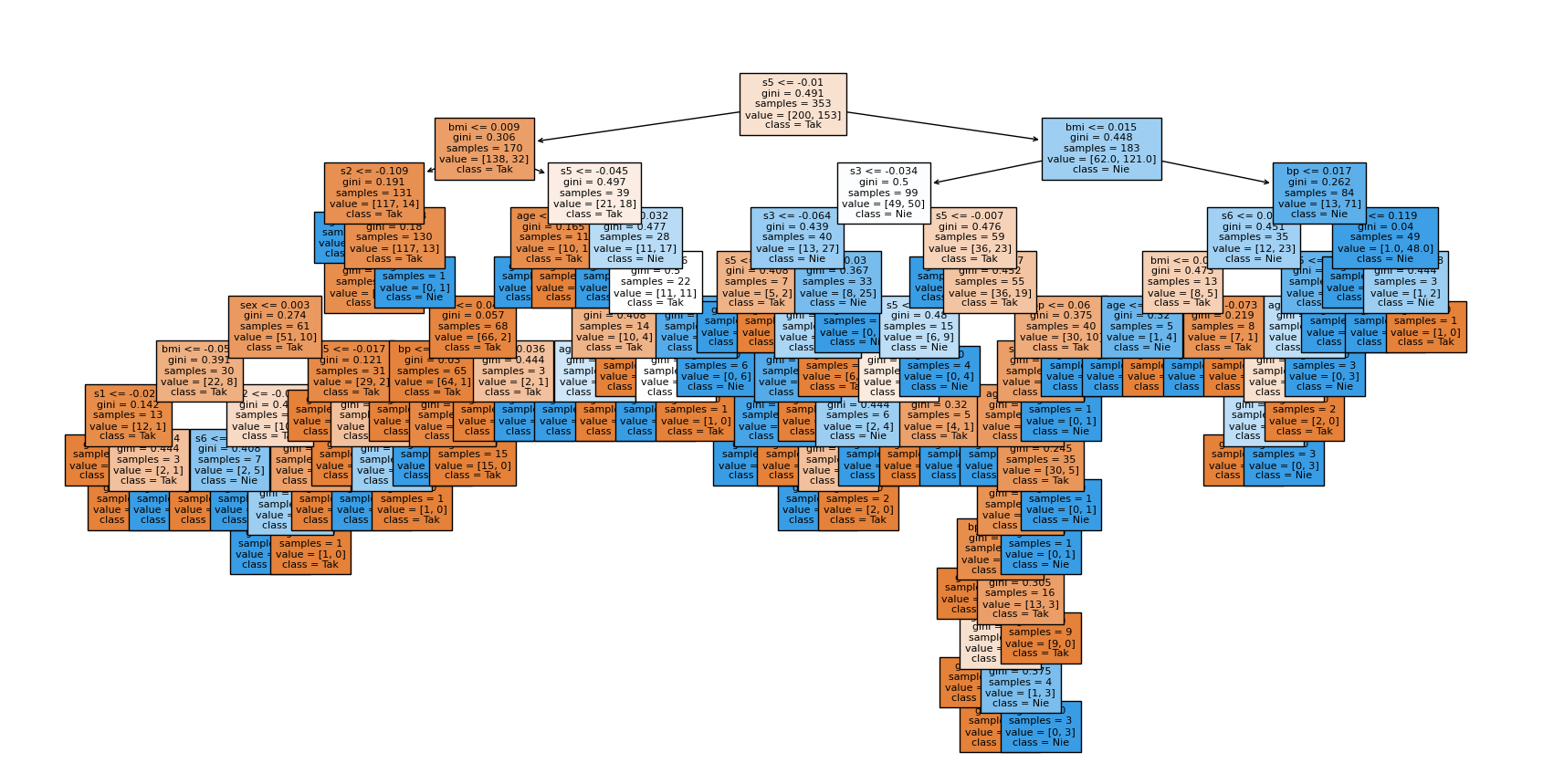
Podział przestrzeni wielowymiarowej na dwie części jest dokładnie tym czego potrzebujemy do naszego zadania. Jest to więc nic dziwnego że model ten radzi sobie tak dobrze.

Model **kNN** również miał dobre wyniki. Jest on jednak nieprzystosowany do użyć wielowymiarowych. Najlepiej radzi on sobie w sytuacji gdzie odległość dwóch punktów ma jakąś sensowną interpretacje. Warto wspomnieć o tym, że parametr k=4 dla modelu **kNN** jest nieoptymalny. Dobór odpowiedniego “k” jest bardzo ważny do odpowiedniego działania modelu.

Odnośnie modelu drzewa decyzyjnego z moich obserwacji wynika iż dochodzi do przetrenowania.

Wiele z liści drzewa ma tylko jeden pasujący element. Jest to jedna z głównych wad tego modelu.

Bardzo łatwo dochodzi do przetrenowania.

 3. Drzewo decyzyjne o domyślnych parametrach