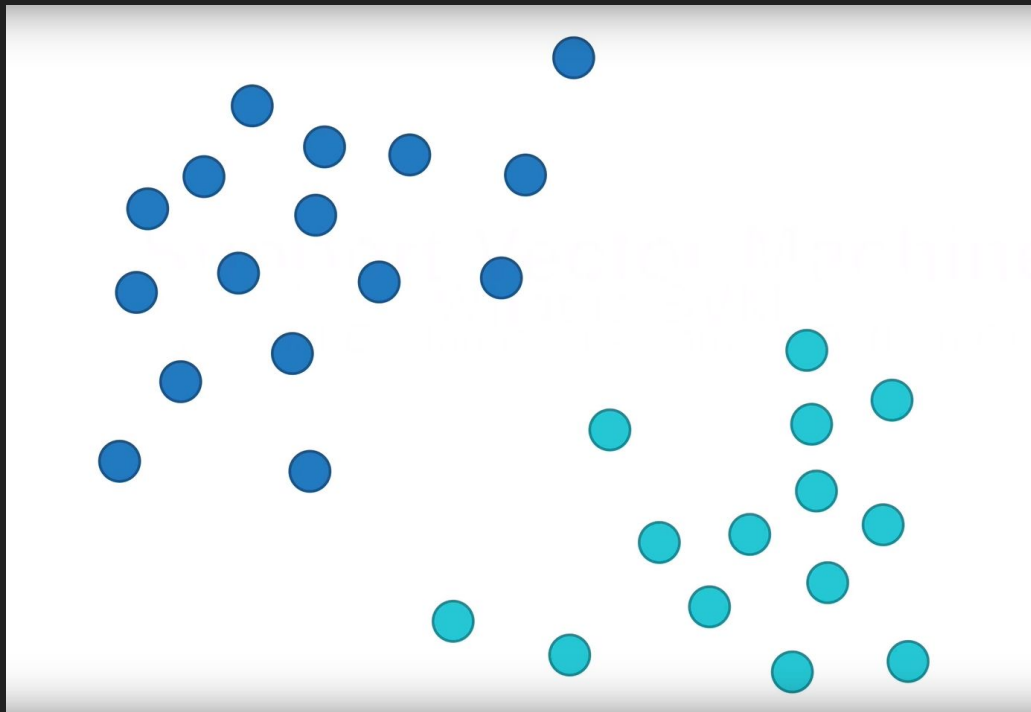
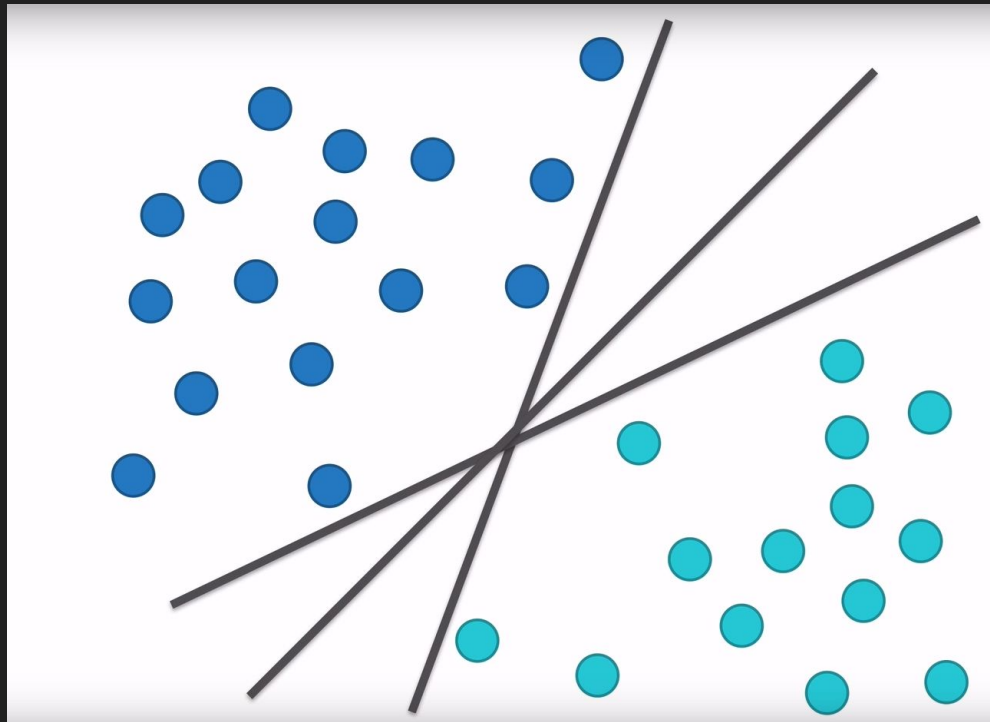


Klasyfikacja i regresja przy pomocy metod SVM

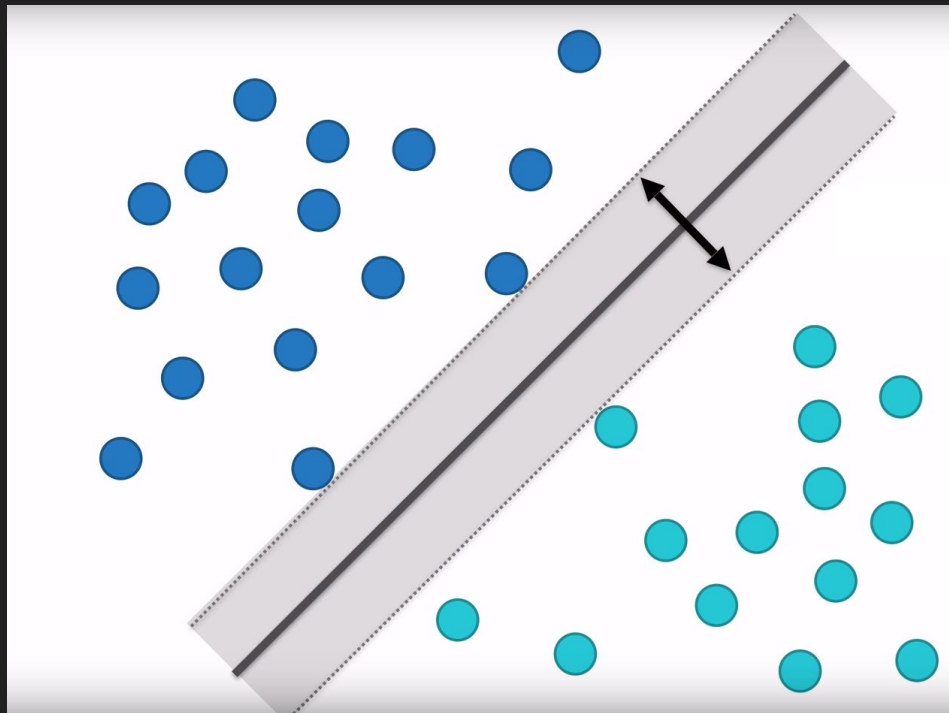
Problem wyznaczania granicy decyzyjnej



Problem wyznaczania granicy decyzyjnej

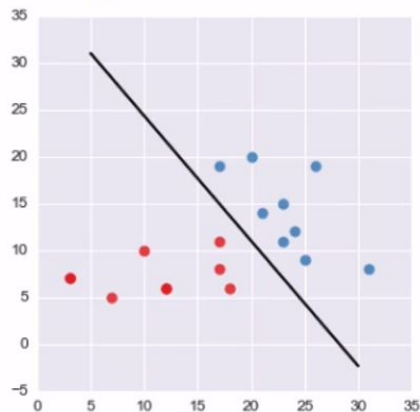


Problem wyznaczania granicy decyzyjnej

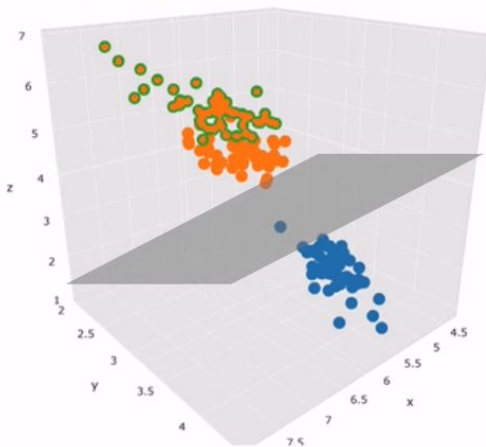


Hiperpłaszczyzna jako granica decyzyjna

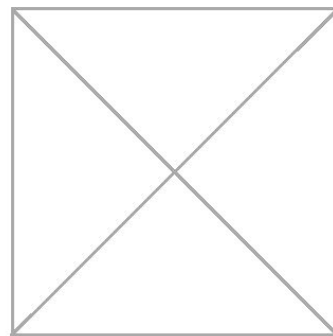
2D: Granicą decyzyjną jest prosta



3D: Granicą decyzyjną jest płaszczyzna

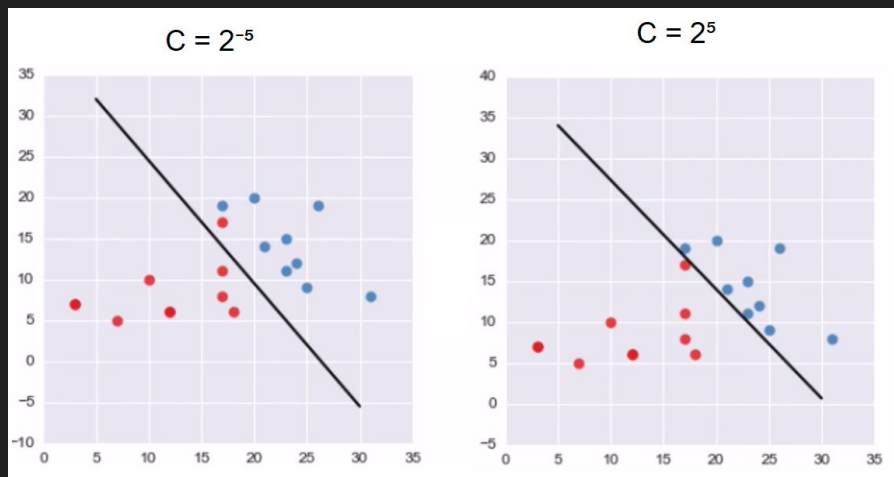


4D+: Granicą decyzyjną jest hiperpłaszczyzna

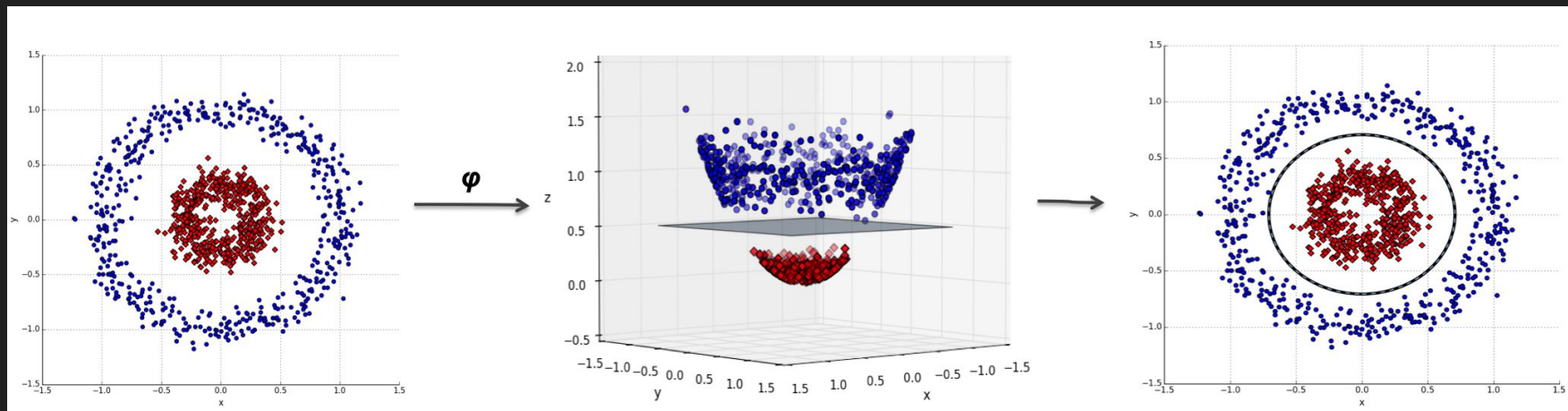


Parametr C

- pozwala decydować w jakim stopniu chcemy unikać błędów klasyfikacji
- mniejsze C = większy margines, więcej błędów klasyfikacji
- większe C = mniejszy margines, mniej błędów klasyfikacji



Kernel tricks



Kernel tricks

Dodanie wymiaru przy pomocy funkcji przekształcającej:

$(a, b) \rightarrow (a, b, \phi(a, b))$

Dostępne przekształcenia:

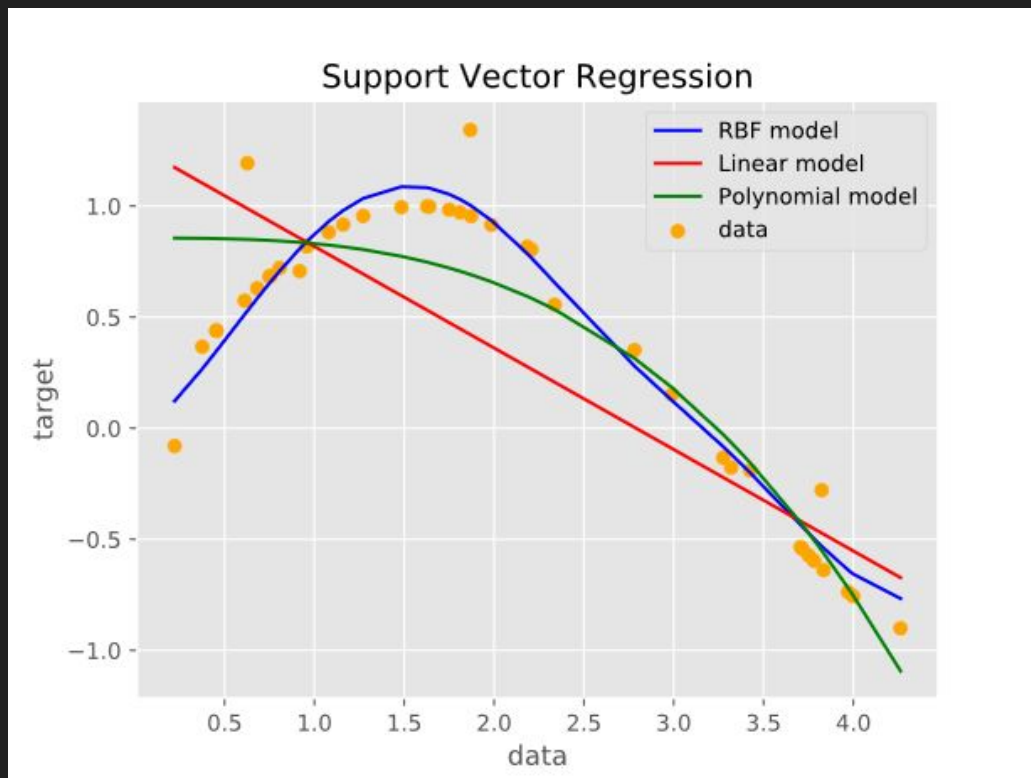
- linear
- rbf
- poly
- sigmoid

Multiclass SVM

Sprowadzenie problemu klasyfikacji wieloklasowej do wielu problemów klasyfikacji binarnej. Dwie popularne metody:

- One vs All: Osobny klasyfikator dla każdej klasy, wygrywa ten o najwyższym wyniku funkcji wyjściowej.
- One vs One: Klasyfikator dla każdej pary klas, wygrywa ta, która wygra w największej liczbie par

Regresja



Krzywa ROC (Receiver Operator Curve)

- Sposób wizualizacji wydajności klasyfikatora binarnego
- Wykres poprawnie zaklasyfikowanych pozytywów do tych, które zostały zaklasyfikowane jako pozytywy nieprawidłowo
- Im większe pole pod krzywą, tym prawdopodobnie lepszy klasyfikator. Linia prosta pod kątem 45 stopni, o jednakowej liczbie pozytywów prawidłowych i fałszywych, odpowiada zupełnie losowej klasyfikacji.

Krzywa ROC (Receiver Operator Curve)

