1**. Podstawowe Informacje**

Support Vector Classifier (SVC) to implementacja algorytmu Support Vector Machine przeznaczona do zadań klasyfikacji. SVC znajduje optymalną hiperpłaszczyznę, która maksymalizuje margines między klasami, wykorzystując tylko najbliższe punkty danych zwane wektorami nośnymi.

Kluczowe Koncepcje

* Support Vectors: Punkty danych najbliższe hiperpłaszczyźnie, które definiują granicę decyzyjną
* Margin: Odległość między hiperpłaszczyzną a najbliższymi punktami każdej klasy
* Kernel Trick: Technika umożliwiająca mapowanie danych do wyższego wymiaru bez jawnej transformacji

3**. Hiperparametry Kluczowe**

**Parametr C** (Regularyzacja)

Parametr C kontroluje trade-off między maksymalizacją marginesu a minimalizacją błędów klasyfikacji:

* Wysokie C (np. 100, 1000): Twardy margines, mała tolerancja na błędy, ryzyko overfittingu
* Niskie C (np. 0.01, 0.1): Miękki margines, większa tolerancja na błędy, może prowadzić do underfittingu
* Typowe zakresy: [0.001, 0.01, 0.1, 1, 10, 100, 1000]

**Parametr Gamma** (dla kerneli nieliniowych)

* Parametr gamma w kontekście Support Vector Classifier (SVC) odpowiada za tzw. "zasięg wpływu" pojedynczego przykładu treningowego na ostateczną granicę decyzyjną modelu. Innymi słowy, gamma określa, jak daleko od danego punktu treningowego jego wpływ na kształtowanie decyzji sięga w przestrzeni cech.
* Wysoka wartość gamma powoduje, że wpływ punktu treningowego jest bardzo lokalny, co prowadzi do bardziej złożonych, szczegółowych granic decyzyjnych i ryzyka przeuczenia (overfittingu).
* Niska wartość gamma daje bardziej rozmyte granice, gdzie każdy punkt ma wpływ na większy obszar, co może skutkować prostszymi granicami i potencjalnym niedouczeniem (underfittingiem).
* Technicznie, w przypadku popularnego kernela RBF (Radial Basis Function), gamma jest parametrem w funkcji jądra definiującym, jak szybko maleje wpływ punktu odległego od centrum:
* Formuła: γ = 1/(2σ²) dla kernela RBF
* Wysokie γ: Złożone granice decyzyjne, każdy punkt ma mały zasięg wpływu
* Niskie γ: Proste granice decyzyjne, każdy punkt ma duży zasięg wpływu
* Wartości domyślne:
  + 'scale': 1/(n\_features \* X.var())
  + 'auto': 1/n\_features

4. Funkcje Kernelowe

Linear Kernel

K(x,y) = x·y

* Zastosowania: Dane liniowo rozdzielne, wysokowymiarowe dane tekstowe
* Zalety: Szybkość obliczeń, interpretowalność

RBF (Radial Basis Function) Kernel

K(x,y) = exp(-γ||x-y||²)

* Najpopularniejszy kernel w praktyce
* Mapuje dane do nieskończonego wymiaru
* Uniwersalny - działa dobrze dla większości problemów

Polynomial Kernel

K(x,y) = (γx·y + r)ᵈ

* Parametry: stopień d , γ (gamma), r (coef0)
* Zastosowania: Natural Language Processing, dane z interakcjami cech
* Uwaga: Może być numerycznie niestabilny dla wysokich stopni

Sigmoid Kernel

K(x,y) = tanh(γx·y + r)

* Charakterystyka: Podobny do sieci neuronowych
* Uwaga: Może nie być pozytywnie półokreślony