

**Data Science
Academy**

www.datascienceacademy.com.br

Machine Learning

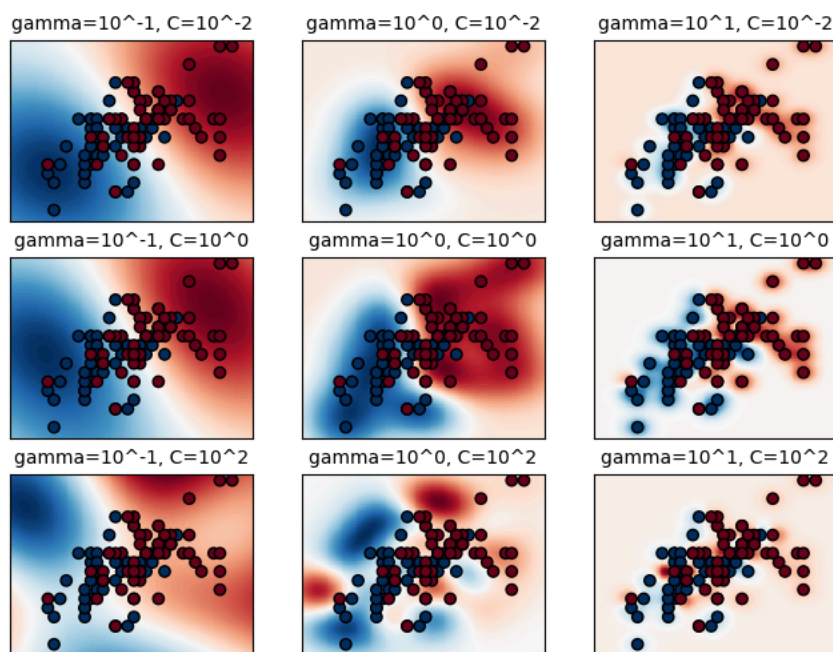
Parâmetros C, Gamma e Kernel RBF

O Kernel RBF é o mais amplamente usado ao treinar modelos SVM e dois Hiperparâmetros nos ajudam a ajustar o modelo ideal: C e gama.

Intuitivamente, o parâmetro gama define até que ponto a influência de um único exemplo de treinamento alcança, com valores baixos significando 'longe' e valores altos significando 'próximos'. Os parâmetros gama podem ser vistos como o inverso do raio de influência das amostras selecionadas pelo modelo como vetores de suporte.

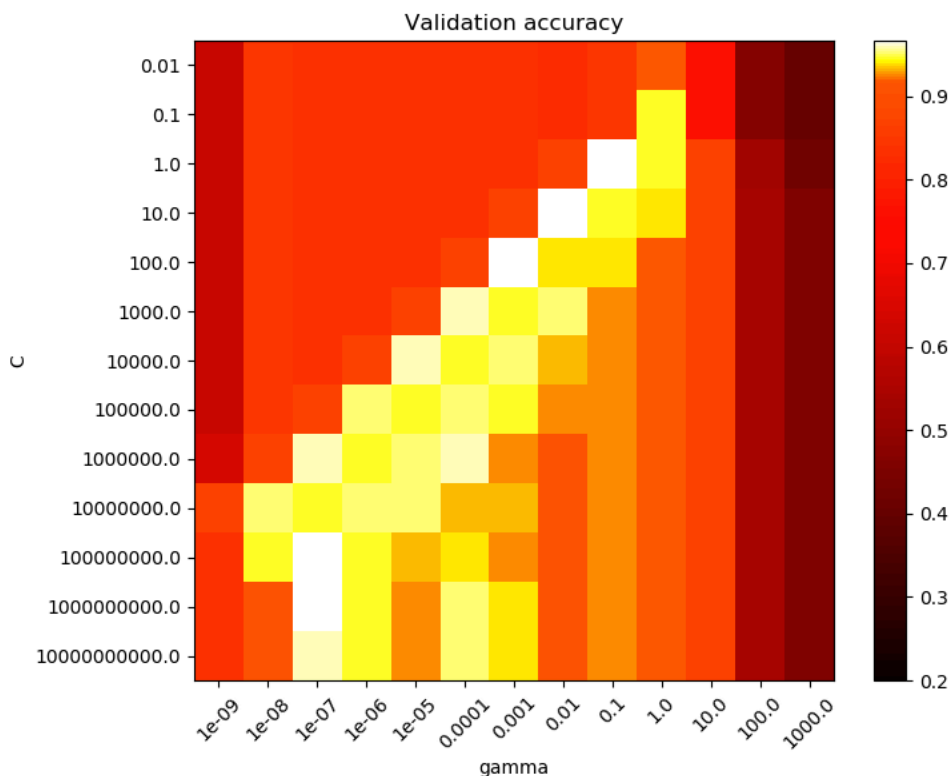
O parâmetro C negocia a classificação correta dos exemplos de treinamento contra a maximização da margem da função de decisão. Para valores maiores de C, uma margem menor será aceita se a função de decisão for melhor na classificação correta de todos os pontos de treinamento. Um C mais baixo incentivará uma margem maior, portanto, uma função de decisão mais simples, ao custo da precisão do treinamento. Em outras palavras, “C” se comporta como um parâmetro de regularização no SVM.

Observe os gráficos abaixo extraídos da documentação do SVM no Scikit-Learn:



O primeiro gráfico é uma visualização da função de decisão para uma variedade de valores de parâmetros em um problema de classificação simplificado, envolvendo apenas 2 recursos de entrada e 2 classes-alvo possíveis (classificação binária). Observe que esse tipo de gráfico não é possível para problemas com mais recursos ou classes de destino.

O segundo gráfico abaixo é um mapa de calor da precisão da validação cruzada do classificador em função de C e γ . Neste exemplo, exploramos uma grade relativamente grande para fins de ilustração. Na prática, uma grade logarítmica geralmente é suficiente. Se os melhores parâmetros estiverem nos limites da grade, eles poderão ser estendidos nessa direção em uma pesquisa subsequente.



Observe que o gráfico do mapa de calor possui uma barra de cores especial com um valor de ponto médio próximo aos valores de pontuação dos modelos com melhor desempenho, para facilitar a diferenciação entre eles em um piscar de olhos.

O comportamento do modelo é muito sensível ao parâmetro γ . Se γ é muito grande, o raio da área de influência dos vetores de suporte inclui apenas o próprio vetor de suporte e nenhuma quantidade de regularização com C será capaz de impedir o ajuste excessivo.

Quando o valor de γ é muito pequeno, o modelo fica muito restrito e não pode capturar a complexidade ou a "forma" dos dados. A região de influência de qualquer vetor de suporte selecionado incluiria todo o conjunto de treinamento. O modelo resultante se comportará de maneira semelhante a um modelo linear com um conjunto de hiperplanos que separam os centros de alta densidade de qualquer par de duas classes.



Para valores intermediários, podemos ver no segundo gráfico que bons modelos podem ser encontrados na diagonal de C e γ . Modelos suaves (valores γ mais baixos) podem ser mais complexos, aumentando a importância de classificar cada ponto corretamente (valores C maiores), daí a diagonal dos modelos com bom desempenho.

Fonte:

https://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html