10/12/2025

MACHINE LEARNING AND NEURAL NETWORKS

7PAM2021-0901-2025

TUTOR – PETER SCICLUNA

Riyanka Dhar [Student-PECS]
UNIVERSITY OF HERTFORDSHIRE

Contents

How SVM Kernels Shape the Decision Boundary: A Visual Tutorial

Author: Riyanka Dhar
Student ID: 23096051
Department: MSC. Data Science
Submission Format: PDF
Submission Date: 10/12/2025

Introduction:

Support Vector Machines (SVMs) are widely used margin-based classifiers that perform effectively on a range of supervised learning tasks. A linear SVM can separate data that is linearly separable in its original feature space, but many real datasets contain non-linear patterns that cannot be represented by a straight decision boundary. Kernel methods address this limitation by allowing SVMs to construct non-linear boundaries without explicitly transforming the input into a higher dimensional space.

This tutorial examines how three commonly used kernels (Linear, Polynomial, and Radial Basis Function) influence the decision boundary on a two-dimensional synthetic dataset. By comparing their behaviour and studying the effect of key hyperparameters such as C and gamma, the analysis aims to provide a clear and visual explanation of how kernel choice affects model flexibility and classification performance.

Background on Support Vector Machines:

Support Vector Machines are supervised learning models used for binary classification and regression tasks. The central idea of an SVM classifier is to identify a separating hyperplane that maximises the margin between the two classes. The margin refers to the distance between the hyperplane and the closest data points from each class, known as support vectors. A larger margin generally leads to better generalisation, since the classifier becomes less sensitive to small variations in the training data.

In the simplest case, the separating hyperplane is linear. However, many datasets are not linearly separable in their original feature space. To address this, SVMs can incorporate a soft margin by introducing a regularisation parameter, C. This parameter controls the balance between maximising the margin and reducing classification errors on the training set. A small value of C encourages a wider margin but may allow some misclassifications, while a large value of C prioritises correct classification, often at the cost of a narrower margin.

The Kernel Method :

Kernel methods extend the flexibility of Support Vector Machines by allowing the classifier to operate in a higher dimensional feature space without explicitly computing new features. Instead of transforming the data through a direct mapping, the SVM relies on a kernel function that measures similarity between data points in the implicit feature space. This approach is known as the kernel trick.

The idea behind the kernel trick is that many linear algorithms become capable of representing complex non-linear patterns when they are applied in a higher dimensional space. With a suitable kernel function, the SVM constructs a non-linear decision boundary while still solving an optimisation problem that retains the structure of the linear case.

Several kernel functions are commonly used in practice. The linear kernel corresponds to the standard dot product and is suitable when the data is linearly separable. The polynomial kernel introduces interactions between features and increases expressive power based on the chosen degree. The Radial Basis Function kernel is highly flexible and can model intricate local patterns by controlling the influence of individual data points through the gamma parameter.

Kernel methods allow SVMs to adapt to a wide variety of classification tasks, making them suitable for both simple and complex datasets.

Dataset and Experimental Setup:

The experiments in this tutorial use a synthetic two-dimensional dataset generated with the make_moons function from the scikit-learn library. This dataset is widely used for illustrating non-linear classification problems because the two classes form interlocking shapes that cannot be separated by a straight line. A small amount of noise is added to the data to produce a more realistic and challenging classification task.

To examine how kernel choice affects the decision boundary, three SVM models are trained using the linear, polynomial, and Radial Basis Function kernels. In addition to the standard experiments, a custom two spiral dataset is generated to demonstrate how kernel methods perform on a more complex pattern. This second dataset highlights the flexibility of the RBF kernel when modelling curved and intertwined structures.

All experiments are carried out in Python using scikit-learn. The notebook associated with this tutorial contains the full implementation, including the code used to train the models, create the decision boundary plots, and export all figures. The parameters C and gamma are varied in separate experiments to illustrate how regularisation and kernel width influence the complexity of the resulting classifier.

Results and Analysis:

This section presents the results of the experiments conducted using the make_moons dataset and the custom spiral dataset. Each kernel is evaluated in terms of its ability to produce an appropriate decision boundary for the underlying structure of the data. The influence of the regularisation parameter C and the RBF kernel width gamma is also examined.

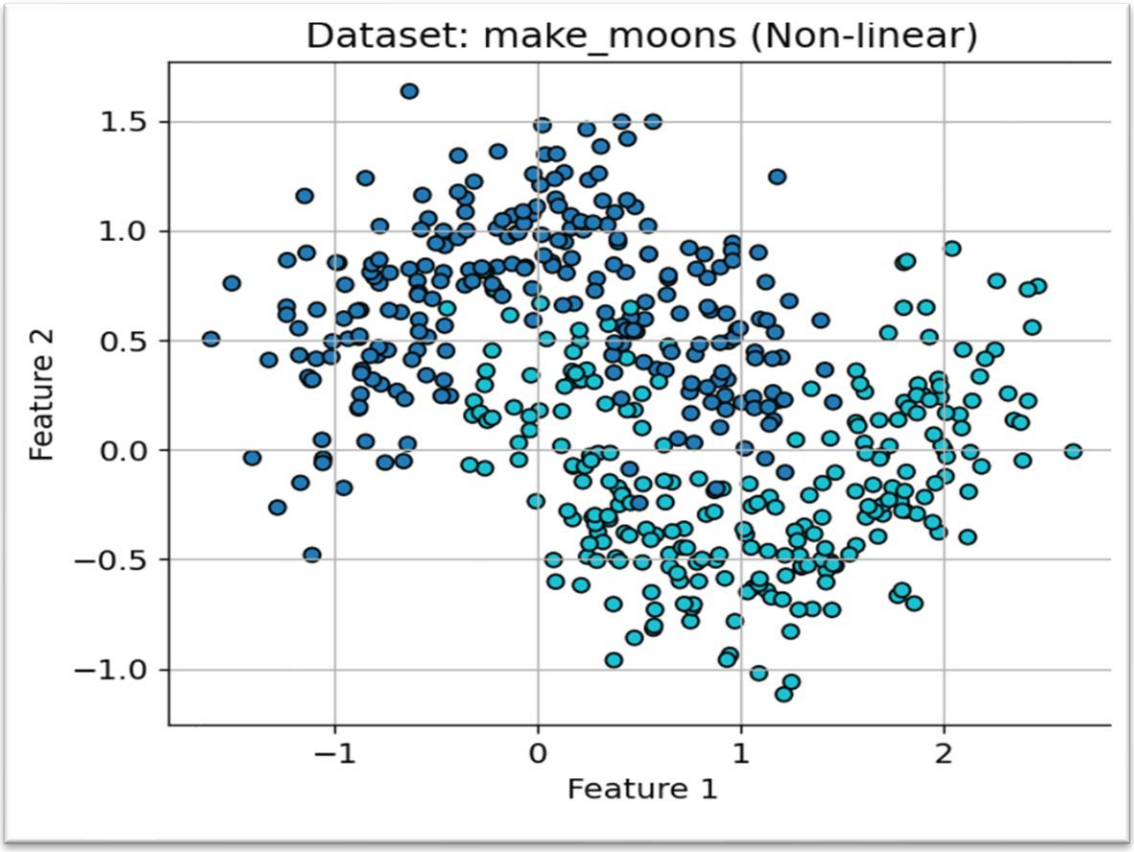

Figure 1. Visualisation of the make_moons dataset with two non-linear classes.

Linear Kernel:

The linear SVM generates a straight decision boundary, which is expected given the form of the kernel. As shown in Figure 1, this boundary does not capture the curved structure of the make_moons dataset and therefore produces a relatively high number of misclassifications near the two class boundaries. This result illustrates the limitation of linear models when applied to non-linear data.

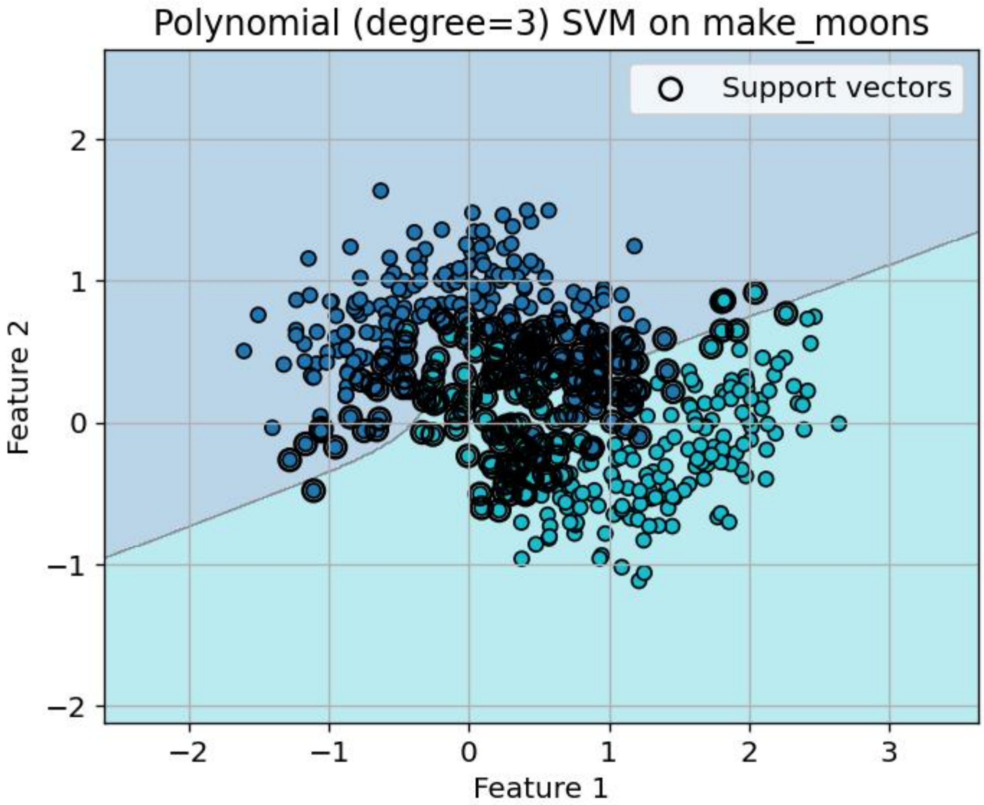

Figure 2. Decision boundary produced by the linear SVM. The straight boundary does not match the curved structure of the dataset.

Polynomial Kernel:

The polynomial SVM introduces curvature into the decision boundary. With degree equal to three, the boundary adapts more closely to the structure of the dataset. Figure 3 shows improved separation compared to the linear kernel. However, higher polynomial degrees may lead to unnecessary complexity and potential overfitting.

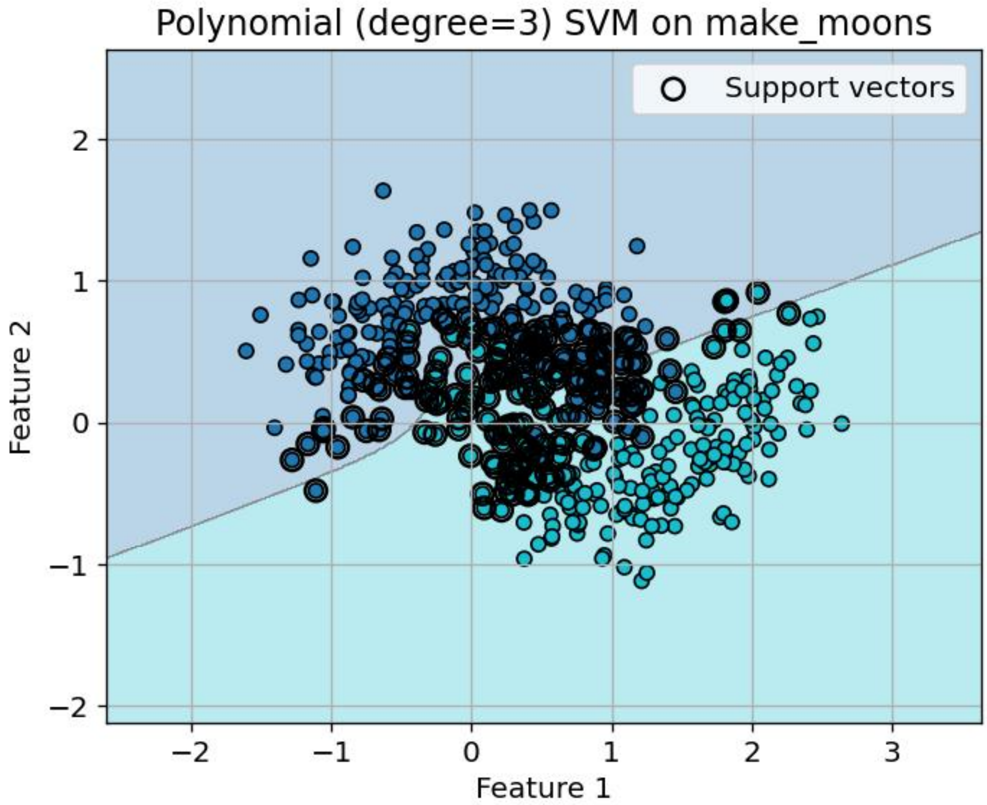

Figure 3. Decision boundary from the polynomial SVM with degree equal to 3.

Radial Basis Function Kernel:

The RBF kernel produces the most flexible decision boundary among the tested models. Figure 4 shows that the boundary follows the curved distribution of the data and separates the classes effectively. The local influence of the RBF kernel allows the classifier to model non-linear relationships with high accuracy.

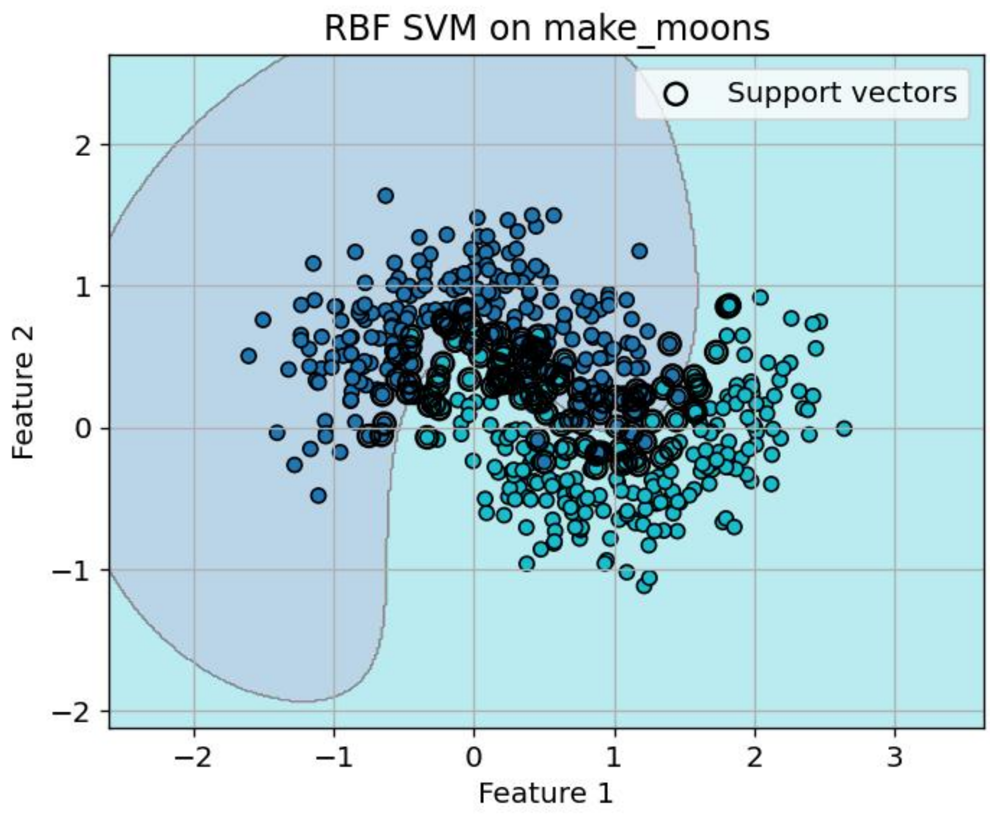

Figure 4. Decision boundary generated by the RBF SVM.

Effect of the Regularisation Parameter C:

The parameter C controls the balance between margin width and misclassification tolerance. A small value of C encourages a wider margin and smoother decision surface. Increasing C forces the model to classify training points more strictly, often producing a tighter and more complex boundary. Figure 5 illustrates this behaviour across three C values.

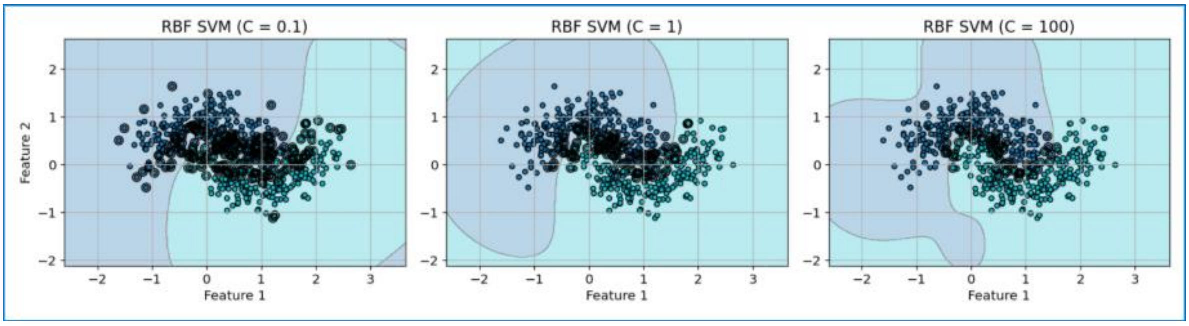

Figure 5. Comparison of RBF SVM decision boundaries for C equal to 0.1, 1, and 100.

Effect of the Gamma Parameter#

Gamma determines how far the influence of a training point extends. A low gamma value results in a smooth, global decision boundary, while a high gamma value produces rapid changes in the decision surface. Figure 6 shows that gamma equal to 10 creates a highly detailed boundary that may overfit the data.

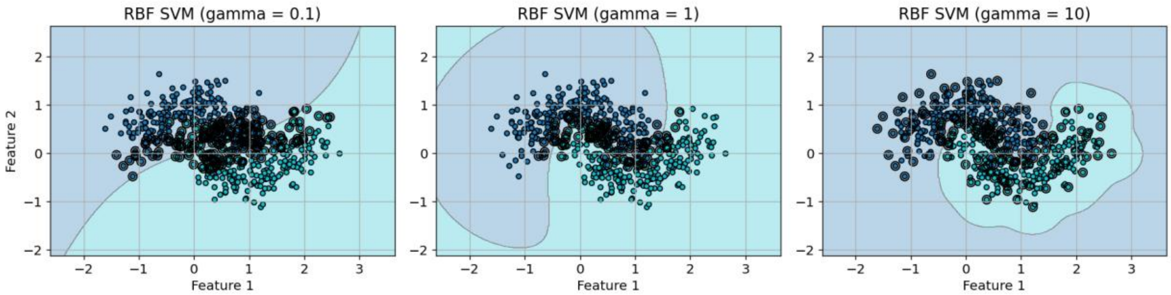

Figure 6. Decision boundaries for the RBF SVM with gamma equal to 0.1, 1, and 10.

Custom Spiral Dataset:

A custom spiral dataset is used to further demonstrate the flexibility of kernel methods. The intertwined structure presents a challenging classification problem. As seen in Figure 7, the RBF SVM can successfully trace the spiral pattern, showing how kernel functions allow SVMs to model highly non-linear structures.

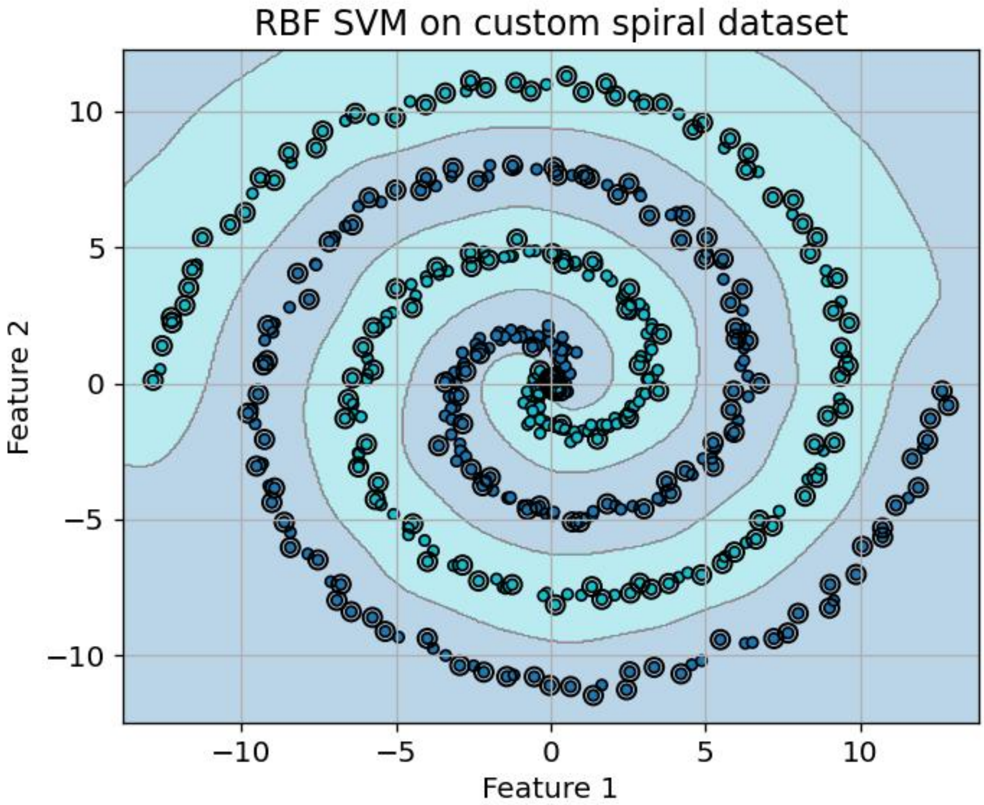

Figure 7. RBF SVM decision boundary on the custom two spiral dataset.

Ethical Considerations:

The use of Support Vector Machines in real applications raises several ethical considerations. Although SVMs are often perceived as relatively transparent compared to some deep learning models, the choice of kernel and hyperparameters can affect the interpretability of the resulting classifier. A linear SVM offers a clear and straightforward explanation of its decision rule, while more complex kernels, such as the polynomial or RBF functions, produce boundaries that are less intuitive to understand. Reduced interpretability may limit the ability of users to assess fairness or identify sources of bias in decision making systems.

The selection and preparation of training data also have ethical implications. If the dataset contains biased or unrepresentative samples, the resulting classifier may reinforce these patterns. This risk is not specific to SVMs but applies to many supervised learning methods. Careful data collection and evaluation are therefore essential when deploying classifiers in sensitive contexts such as finance, healthcare, or employment.

Another consideration relates to transparency. Practitioners should communicate the influence of hyperparameters, such as C and gamma, since these values can affect the sensitivity of the model to individual data points. When a model is used to support decisions that affect individuals, it is important to document how the classifier was trained, validated, and tested.

Overall, ethical use of SVMs requires attention to data quality, model transparency, and the suitability of the chosen kernel for the intended application.

Conclusion:

The experiments presented in this tutorial demonstrate how the choice of kernel influences the behaviour and performance of Support Vector Machine classifiers. The linear kernel provides a simple and interpretable decision boundary, but it is not suitable for datasets with curved or non-linear structure. The polynomial kernel offers additional flexibility by introducing higher order interactions between features. The Radial Basis Function kernel provides the most adaptable decision surface and performs well on both the make_moons dataset and the more complex spiral dataset.

The analysis of the hyperparameters C and gamma further shows how model complexity can be adjusted. Lower values of C promote smoother boundaries through wider margins, while higher values focus on correctly classifying training samples. Similarly, gamma controls how quickly the decision surface responds to local variations in the data. Understanding the influence of these parameters is essential for selecting and tuning SVM models in practice.

Overall, the results highlight the value of kernel methods when working with non-linear datasets and provide a foundation for applying SVMs effectively in a range of classification tasks.

References:

Bishop, C. M., 2006. *Pattern Recognition and Machine Learning*. New York: Springer.

Cortes, C. and Vapnik, V., 1995. Support vector networks. *Machine Learning*, 20(3), pp.273–297.

Goodfellow, I., Bengio, Y. and Courville, A., 2016. *Deep Learning*. Cambridge, MA: MIT Press.

Hastie, T., Tibshirani, R. and Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer.

Pedregosa, F. et al., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, pp.2825–2830.

Scikit-learn, 2025. Support Vector Machines documentation. Available at: https://scikit-learn.org/stable/modules/svm.html (Accessed: 08/12/2025).

GitHub Repository:

The full Jupyter notebook, figures, and supporting code for this tutorial are available in the accompanying GitHub repository: https://github.com/RiyankaDhar/SVM-Kernel-Tutorial.git