SCIENCE ACADEMY
Excellence in Teaching & Learning

Support Vector Machines — SVMs

Hi Guys,
Welcome to the Support Vector Machines lecture!

In this lecture, we will talk about another widely used supervised machine learning technique, Support Vector Machines (SVMs). SVMs deliver state-of-the-art performance in a range of real-world application and considered as one of the standard tool for machine learning. We will explore the theory behind SVMs and try to understand their working principle with practical examples.

✅ *Optional Readings and References:*
sklearn's Official Documentation
Introduction to Statistical Learning - Chapter 9
Machine Learning - A Probabilistic Perspective - Chapter 14

General message: *Key concepts along with significant commentary / text is provided in the slides, so that they serve as a reference for the respective theory lecture. However, the suggested readings are recommended to explore more on the topic under discussion! Good luck!*

Dr. Junaid S. Qazi PhD

Support Vector Machines — SVMs

Support vector machines (SVMs) are a set of supervised learning methods used for both classification and regression analysis. SVMs have been shown to perform well in a variety of settings, and are often considered one of the best “out of the box” classifiers.

Support Vector Machines — SVMs

Support vector machines (SVMs) are a set of supervised learning methods used for both classification and regression analysis. SVMs have been shown to perform well in a variety of settings, and are often considered one of the best "out of the box" classifiers.

- Given a set of training examples, each marked as belonging to one or the other of two categories, SVM training algorithm builds a model that assigns new examples (test data-points) to one category or the other, making it a non-probabilistic binary linear classifier.
- SVM model is a representation of the examples as points in n-dimensional space (n is the number of features), mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible.
- New examples are then mapped into that same n-dimensional space and predictions are made for a category based on which side of the gap the new examples fall.

SCIENCE ACADEMY
Excellence in Teaching & Learning

Support Vector Machines — SVMs

Support vector machines (SVMs) are a set of supervised learning methods used for both classification and regression analysis. SVMs have been shown to perform well in a variety of settings, and are often considered one of the best “out of the box” classifiers.

- Given a set of training examples, each marked as belonging to one or the other of two categories, SVM training algorithm builds a model that assigns new examples (test data-points) to one category or the other, making it a non-probabilistic binary linear classifier.
- SVM model is a representation of the examples as points in n-dimensional space (n is the number of features), mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible.
- New examples are then mapped into that same n-dimensional space and predictions are made for a category based on which side of the gap the new examples fall.

SCIENCE ACADEMY
Excellence in Teaching & Learning

Support Vector Machines — SVMs

Support vector machines (SVMs) are a set of supervised learning methods used for both classification and regression analysis. SVMs have been shown to perform well in a variety of settings, and are often considered one of the best “out of the box” classifiers.

- Given a set of training examples, each marked as belonging to one or the other of two categories, SVM training algorithm builds a model that assigns new examples (test data-points) to one category or the other, making it a non-probabilistic binary linear classifier.
- SVM model is a representation of the examples as points in n-dimensional space (n is the number of features), mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible.
- New examples are then mapped into that same n-dimensional space and predictions are made for a category based on which side of the gap the new examples fall.

Support Vector Machines — SVMs

Support vector machines (SVMs) are a set of supervised learning methods used for both classification and regression analysis. SVMs have been shown to perform well in a variety of settings, and are often considered one of the best “out of the box” classifiers.

- Given a set of training examples, each marked as belonging to one or the other of two categories, SVM training algorithm builds a model that assigns new examples (test data-points) to one category or the other, making it a non-probabilistic binary linear classifier.
- SVM model is a representation of the examples as points in n-dimensional space (n is the number of features), mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible.
- New examples are then mapped into that same n-dimensional space and predictions are made for a category based on which side of the gap the new examples fall.

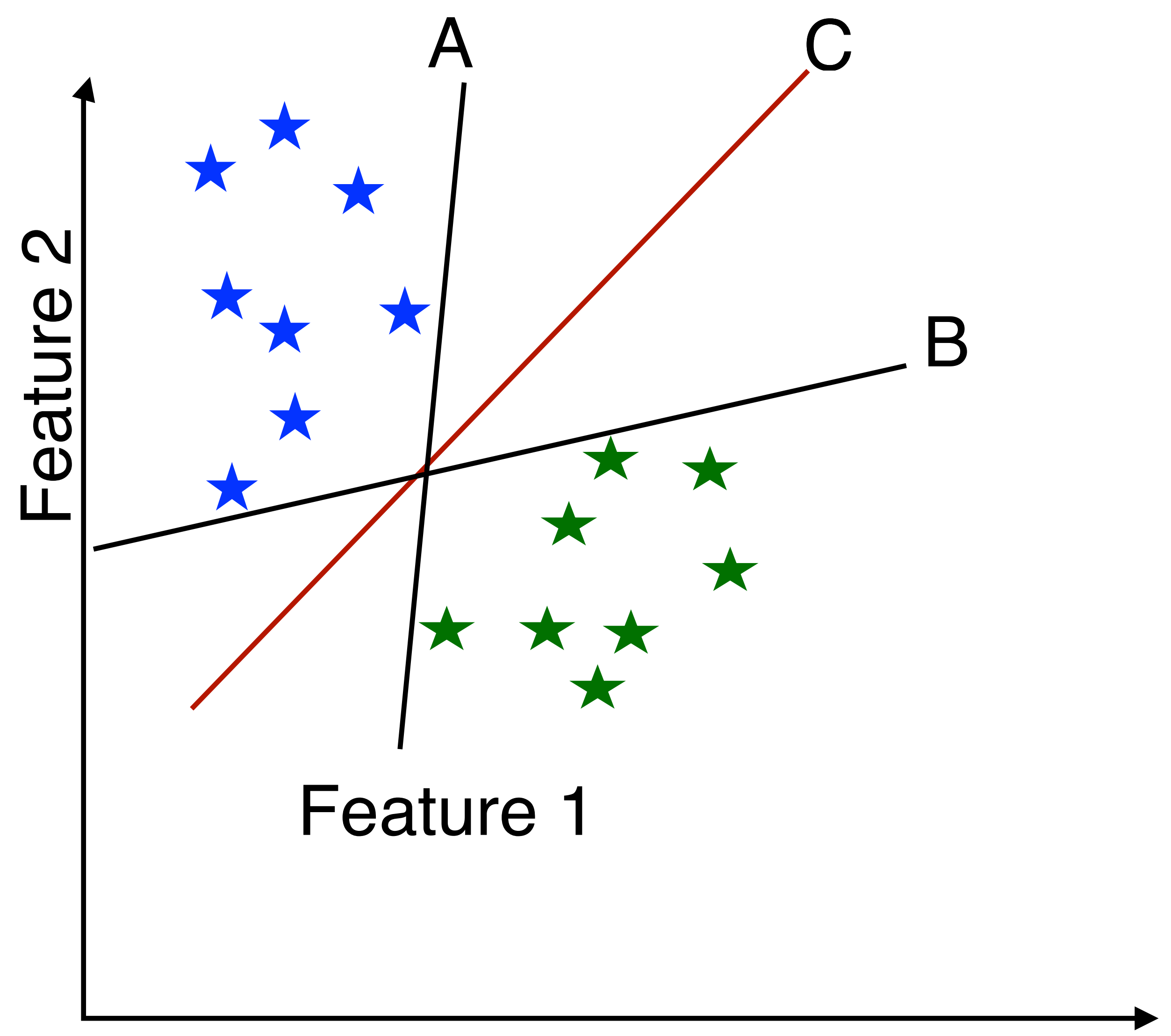
Let’s try to understand with some example training dataset.

Support Vector Machines — SVMs

Consider, we have a labeled training data, blue and green with some feature 1 and 2. We want to predict the class for our new point in this binary classification problem.

Support Vector Machines — SVMs

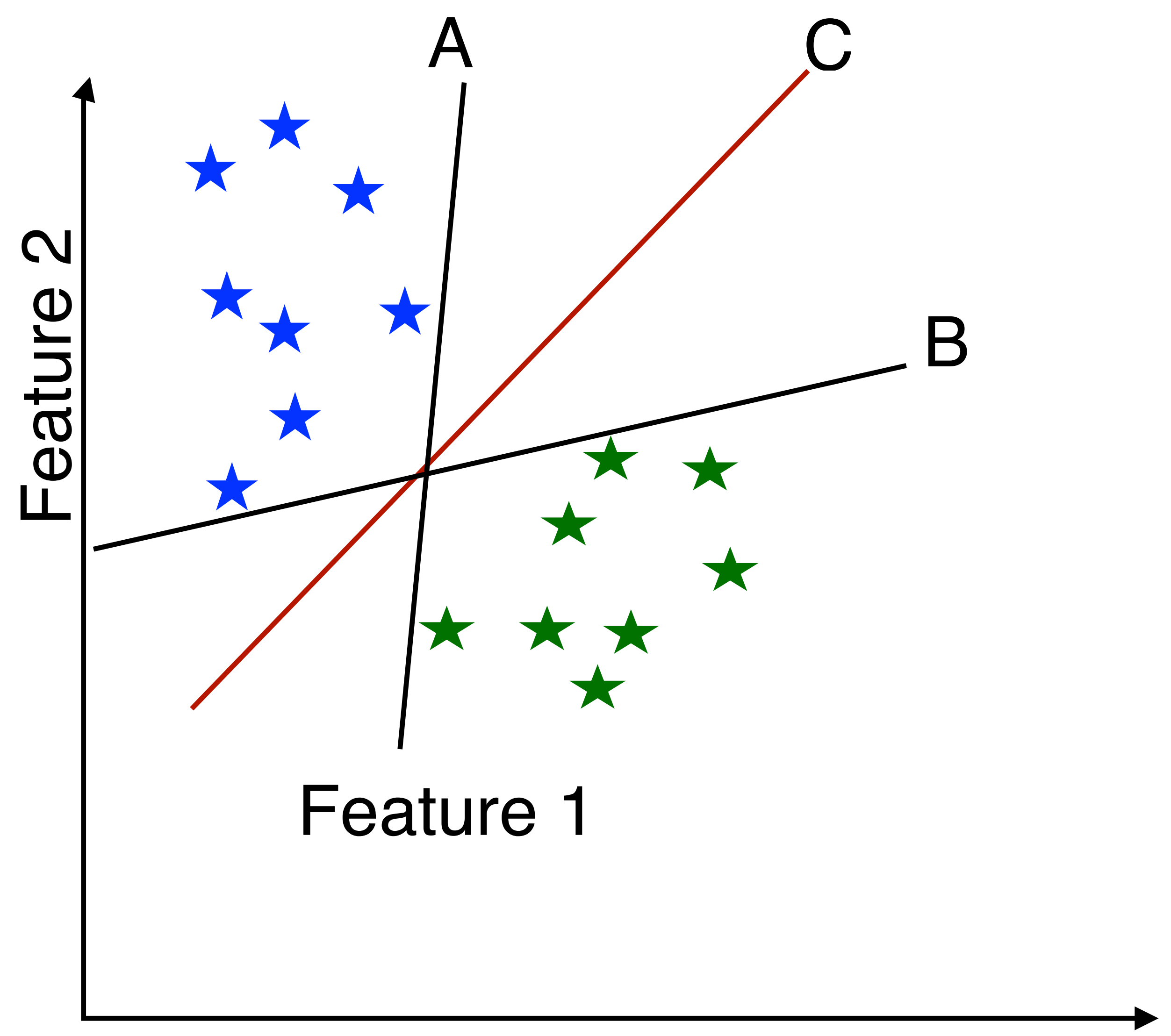
Consider, we have a labeled training data, blue and green with some feature 1 and 2. We want to predict the class for our new point in this binary classification problem.



We draw a hyperplane (simple line in this 2D data) between the two classes. To separate the classes perfectly, we have lots of options such as A, B and C in the diagram (we can draw more hyper planes!). The question is, which plane is the best to select!

Support Vector Machines — SVMs

Consider, we have a labeled training data, blue and green with some feature 1 and 2. We want to predict the class for our new point in this binary classification problem.



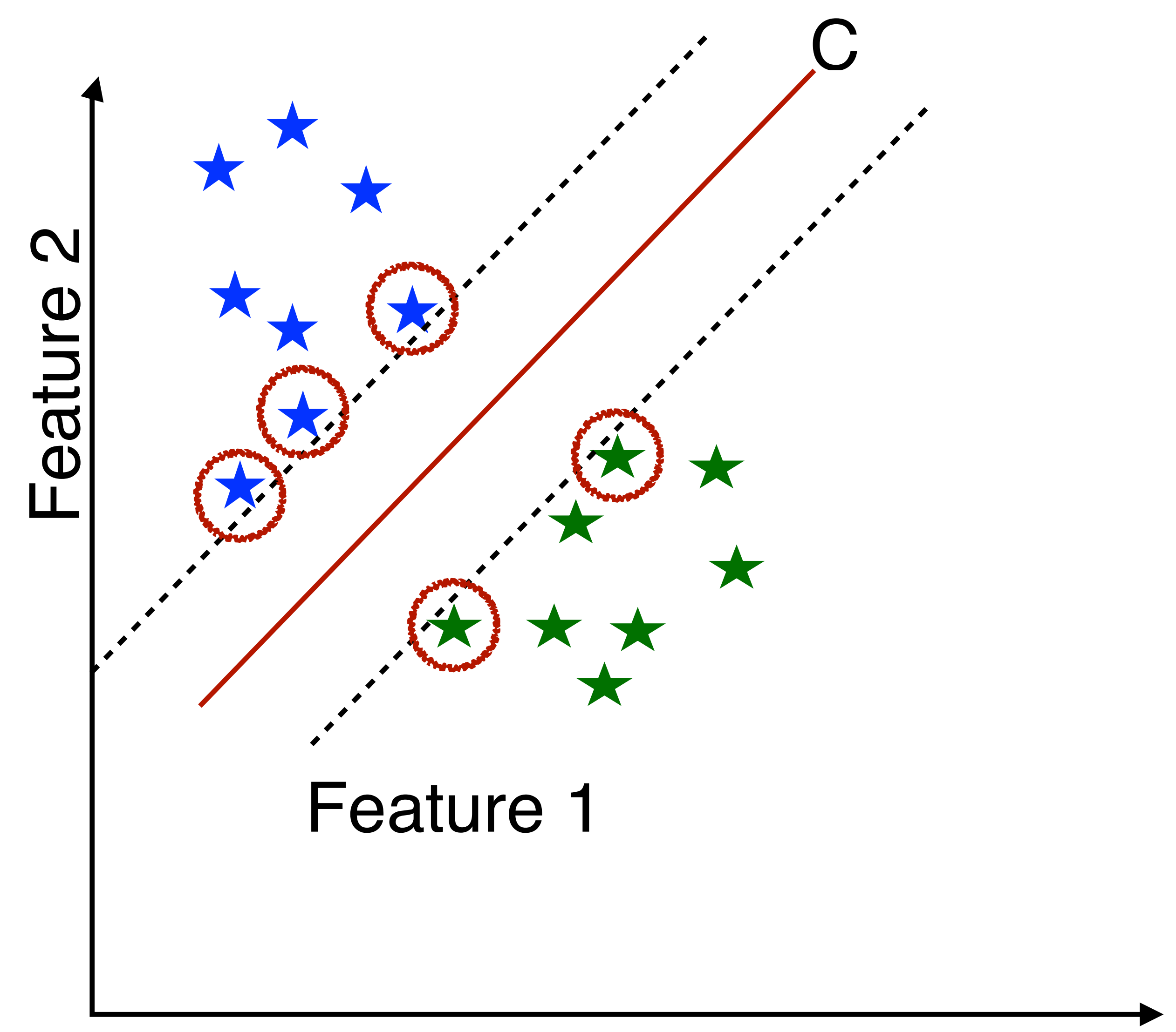
We draw a hyperplane (simple line in this 2D data) between the two classes. To separate the classes perfectly, we have lots of options such as A, B and C in the diagram (we can draw more hyper planes!). The question is, which plane is the best to select!

The plane which maximize the margin between the classes is the best option, in this case C separates the classes with the maximum margin. A & B also separate the classes, but only with a small margin.

SCIENCE ACADEMY
Excellence in Teaching & Learning

Support Vector Machines — SVMs

Consider, we have a labeled training data, blue and green with some feature 1 and 2. We want to predict the class for our new point in this binary classification problem.



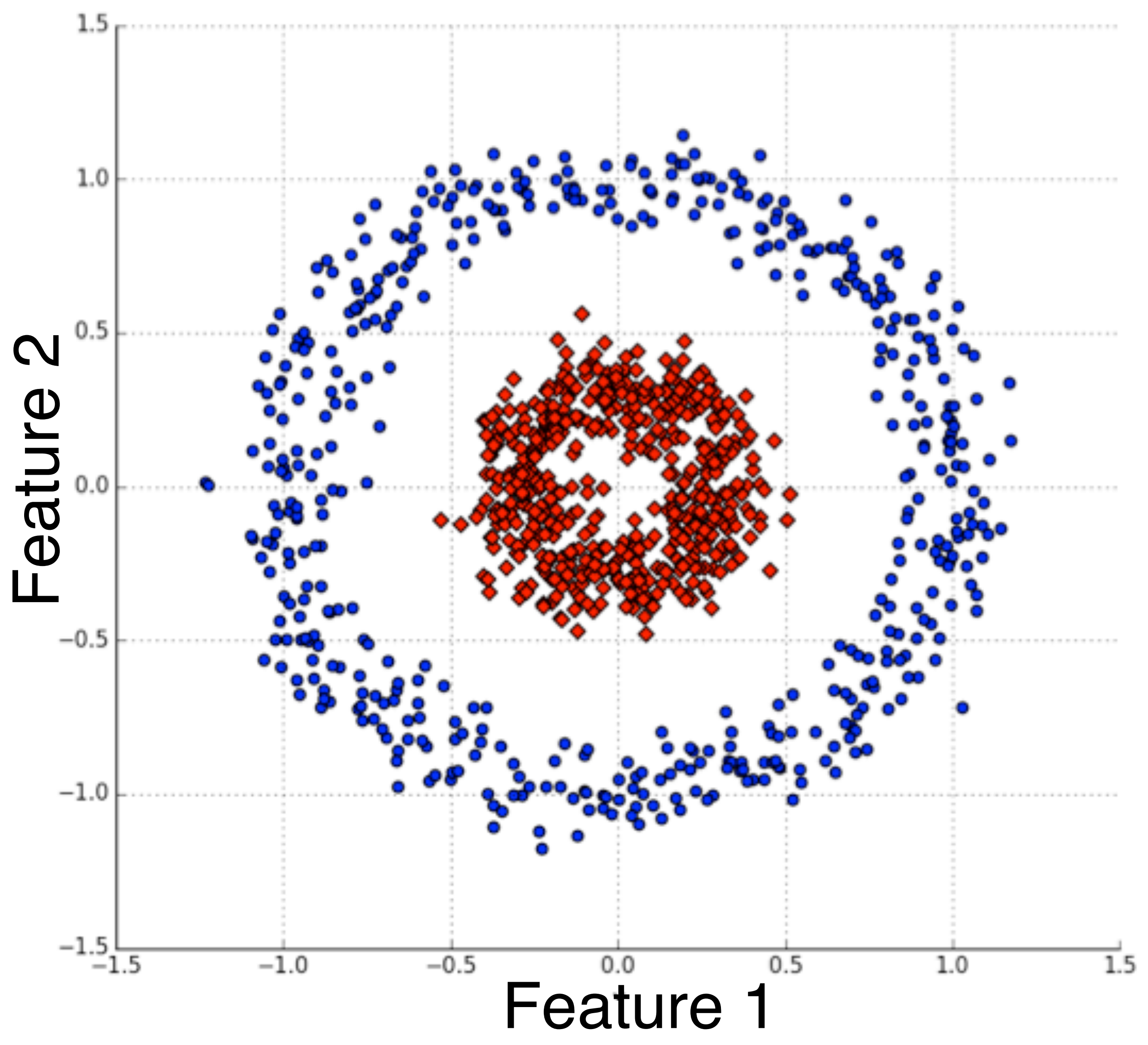
We draw a hyperplane (simple line in this 2D data) between the two classes. To separate the classes perfectly, we have lots of options such as A, B and C in the diagram (we can draw more hyper planes!). The question is, which plane is the best to select!

The plane which maximize the margin between the classes is the best option, in this case C separates the classes with the maximum margin. A & B also separate the classes, but only with a small margin.

In the diagram, dotted lines are the margin lines that extend out from the hyperplane. The vector point (training points), in circles, that are touching the the dotted (margin) lines are called Support Vectors and this is where the name Support Vector Machines comes from.

Support Vector Machines — SVMs

In addition to performing linear classification, SVMs can efficiently perform a **non-linear classification** using what is called the kernel trick, implicitly mapping their inputs into higher-dimensional feature spaces.

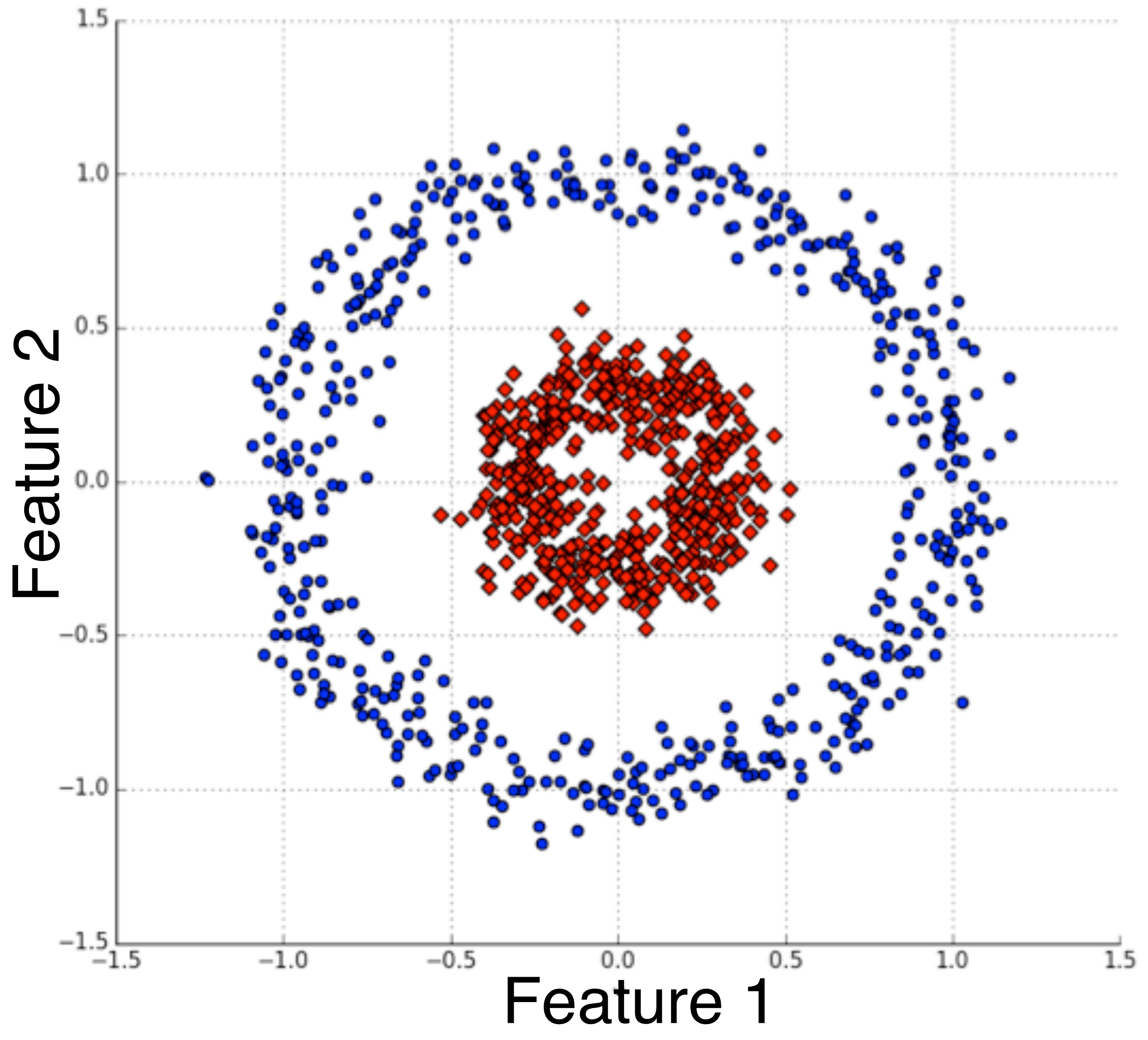


non-linearly separable data, the input space in this case cannot be separated well by a linear classifier and yet we can obviously see a clear separation between the two classes.

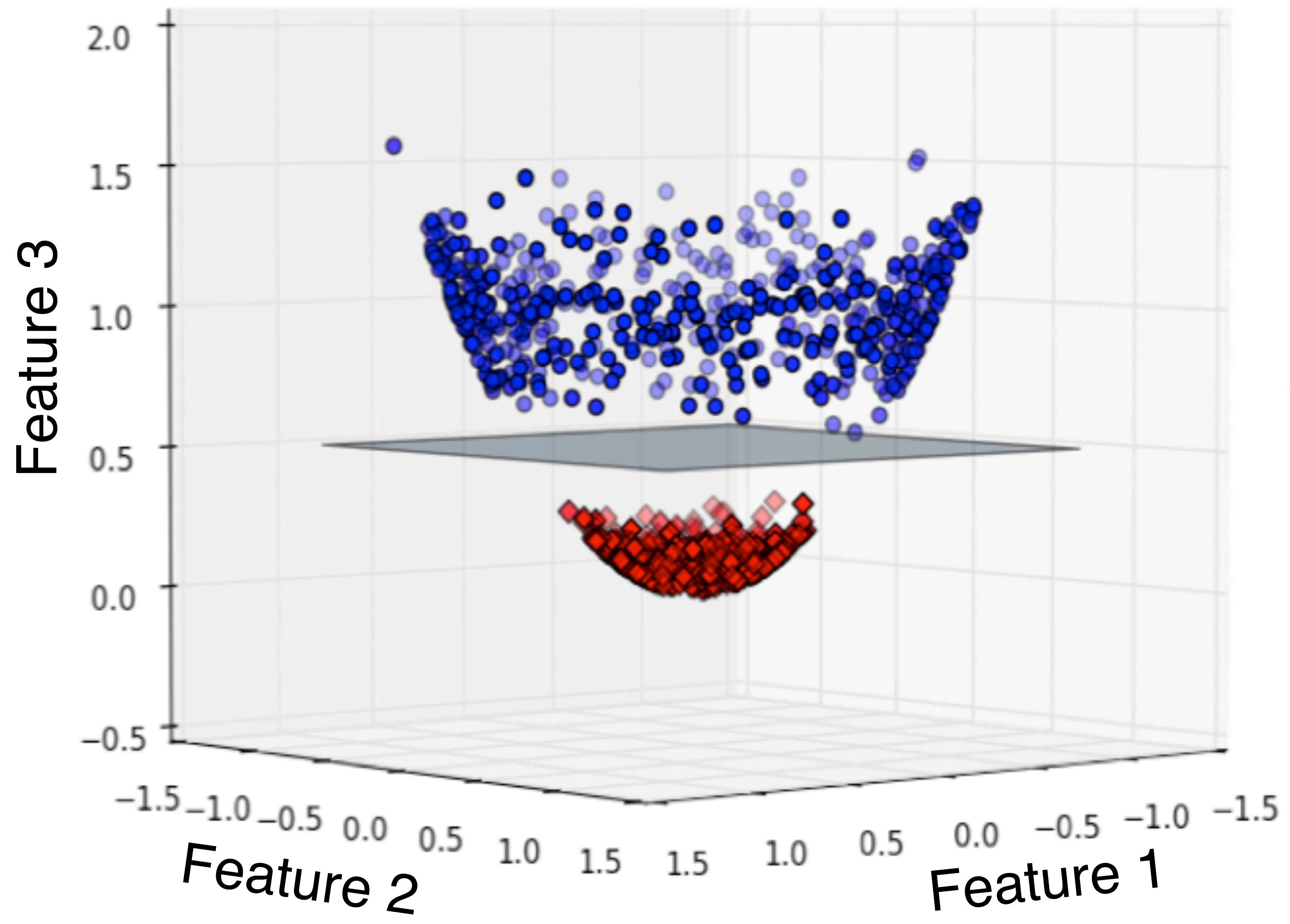
Figure Source, fair policy use

Support Vector Machines — SVMs

In addition to performing linear classification, SVMs can efficiently perform a **non-linear classification** using what is called the kernel trick, implicitly mapping their inputs into higher-dimensional feature spaces.



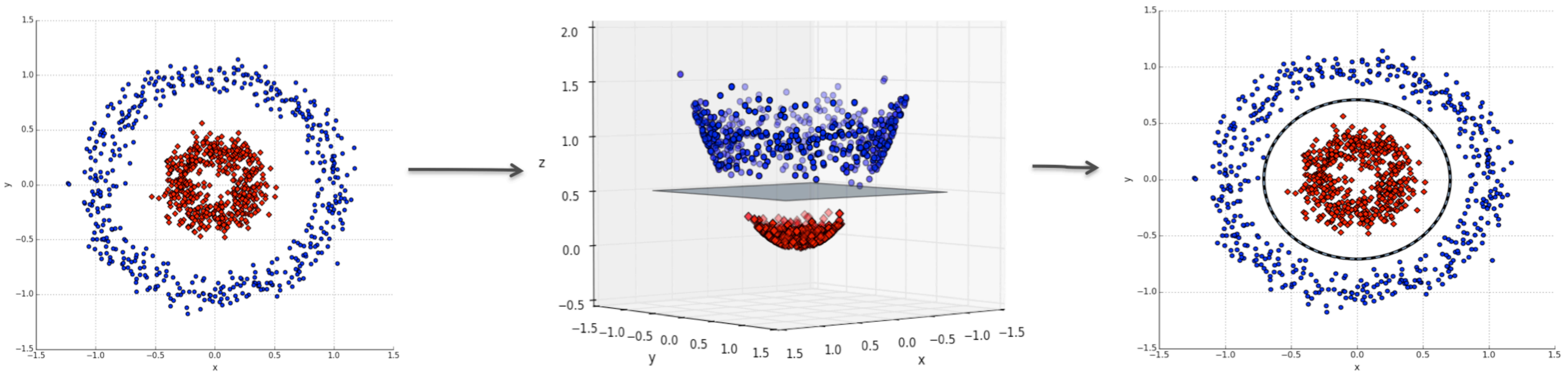
non-linearly separable data, the input space in this case cannot be separated well by a linear classifier and yet we can obviously see a clear separation between the two classes.



The kernel trick (viewing the data in higher dimensional feature space) transforms the data in a feature space where the instances from the two classes may be linearly separable.

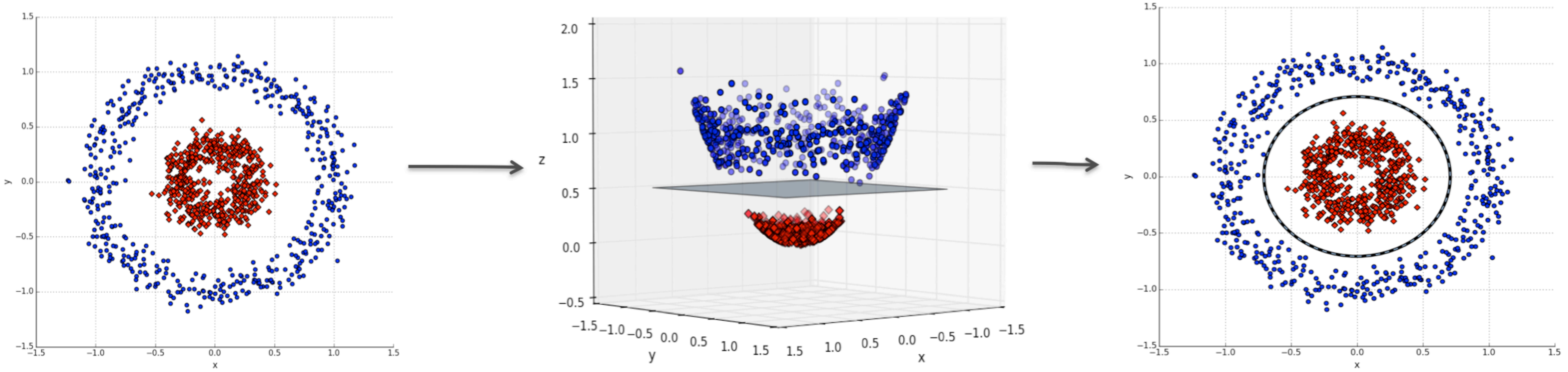
Figure Source, fair policy use

Support Vector Machines — SVMs



So, using the kernel trick, we end up separating the non-linearly separable data by viewing it in higher dimensional feature space.

Support Vector Machines — SVMs



So, using the kernel trick, we end up separating the non-linearly separable data by viewing it in higher dimensional feature space.

So far, we are done with the theory on SVMs. In this lecture we have coved the key points and principle behind SVMs,. Let's learn by doing and see how the things actually work in reality.
In the next lecture, we will explore the Breast Cancer dataset using SVMs. In this way, we will also be able to compare and see the differences between the other algorithms that we have trained and deployed on this data!