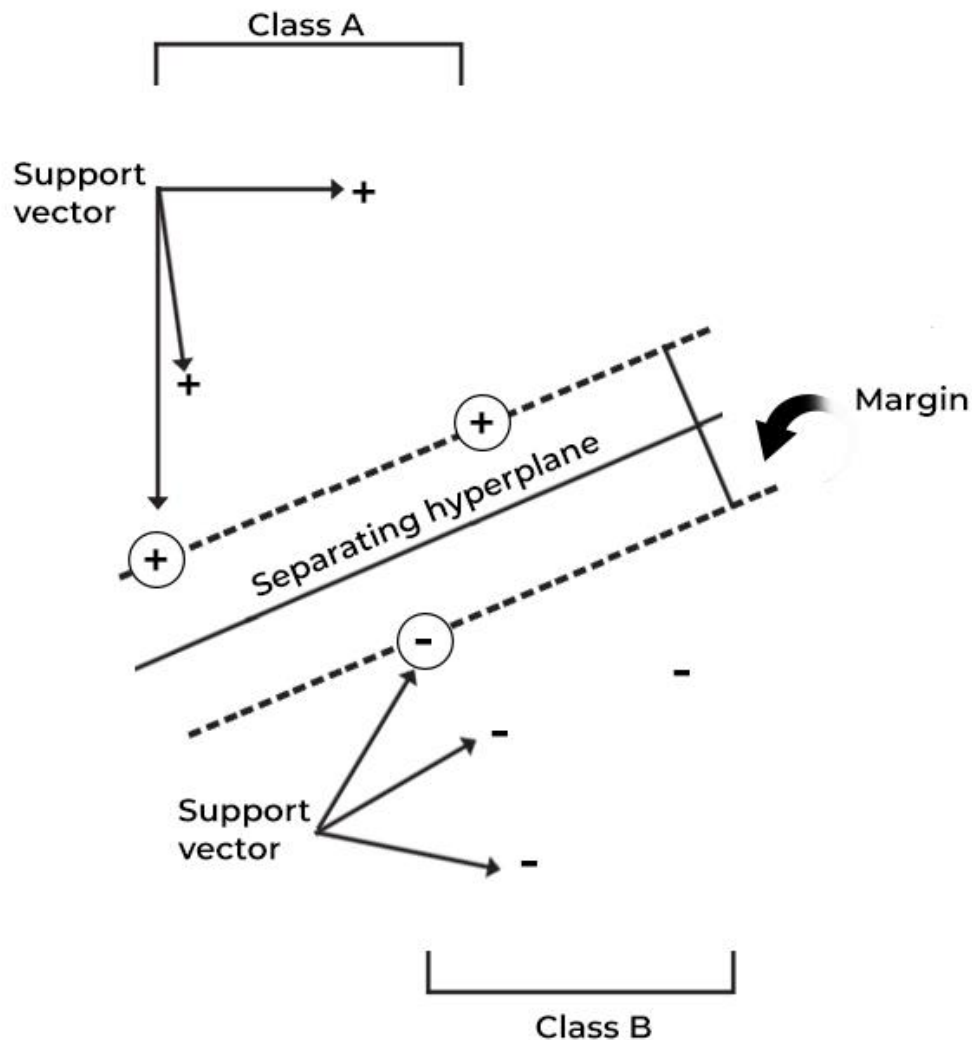


What Is a Support Vector Machine?

A support vector machine (SVM) is a machine learning algorithm that uses supervised learning models to solve complex classification, regression, and outlier detection problems by performing optimal data transformations that determine boundaries between data points based on predefined classes, labels, or outputs. SVMs are widely adopted across disciplines such as healthcare, natural language processing, signal processing applications, and speech & image recognition fields. Technically, the primary objective of the SVM algorithm is to identify a hyperplane that distinguishably segregates the data points of different classes. The hyperplane is localized in such a manner that the largest margin separates the classes under consideration. The support vector representation is shown in the figure below:

SVMS OPTIMIZE MARGIN BETWEEN SUPPORT VECTORS OR CLASSES



SVMs Optimize Margin Between Support Vectors or Classes

As seen in the above figure, the margin refers to the maximum width of the slice that runs parallel to the hyperplane without any internal support vectors. Such hyperplanes are easier to define for linearly separable problems; however, for real-life problems or scenarios, the SVM algorithm tries to maximize the margin between the support

vectors, thereby giving rise to incorrect classifications for smaller sections of data points.

SVMs are potentially designed for binary classification problems. However, with the rise in computationally intensive multiclass problems, several binary classifiers are constructed and combined to formulate SVMs that can implement such multiclass classifications through binary means.

In the mathematical context, an SVM refers to a set of [ML algorithms](#) that use kernel methods to transform data features by employing kernel functions. Kernel functions rely on the process of mapping complex datasets to higher dimensions in a manner that makes data point separation easier. The function simplifies the data boundaries for non-linear problems by adding higher dimensions to map complex data points.

While introducing additional dimensions, the data is not entirely transformed as it can act as a computationally taxing process. This technique is usually referred to as the kernel trick, wherein data transformation into higher dimensions is achieved efficiently and inexpensively.

The idea behind the SVM algorithm was first captured in 1963 by Vladimir N. Vapnik and Alexey Ya. Chervonenkis. Since then, SVMs have gained enough popularity as they have continued to have wide-scale implications across several areas, including the protein sorting process, text categorization, facial recognition, autonomous cars, robotic systems, and so on.

How Does a Support Vector Machine Work?

The working of a support vector machine can be better understood through an example. Let's assume we have red and black labels with the features denoted by x and y . We intend to have a classifier for these tags that classifies data into either the red or black category. Let's plot the labeled data on an x - y plane, as below:

A typical SVM separates these data points into red and black tags using the hyperplane, which is a two-dimensional line in this case. The hyperplane denotes the decision boundary line, wherein data points fall under the red or black category.

A hyperplane is defined as a line that tends to widen the margins between the two closest tags or labels (red and black). The distance of the hyperplane to the most immediate label is the largest, making the data classification easier.

The above scenario is applicable for linearly separable data. However, for non-linear data, a simple straight line cannot separate the distinct data points.

Here's an example of non-linear complex dataset data:

The above dataset reveals that a single hyperplane is not sufficient to separate the involved labels or tags. However, here, the vectors are visibly distinct, making segregating them easier.

For data classification, you need to add another dimension to the feature space. For linear data discussed until this point, two dimensions of x and y were sufficient. In this case, we add a z -dimension to better classify the data points. Moreover, for convenience, let's use the equation for a circle, $z = x^2 + y^2$.

With the third dimension, the slice of feature space along the z -direction looks like this:

Now, with three dimensions, the hyperplane, in this case, runs parallel to the x -direction at a particular value of z ; let's consider it as $z=1$.

The remaining data points are further mapped back to two dimensions.

The above figure reveals the boundary for data points along features x , y , and z along a circle of the circumference with radii of 1 unit that segregates two labels of tags via the SVM.

Let's consider another method of visualizing data points in three dimensions for separating two tags (two different colored tennis balls in this case). Consider the balls lying on a 2D plane surface. Now, if we lift the surface upward, all the tennis balls are distributed in the air. The two differently colored balls may separate in the air at one point in this process. While this occurs, you can use or place the surface between two segregated sets of balls.

In this entire process, the act of 'lifting' the 2D surface refers to the event of mapping data into higher dimensions, which is technically referred to as 'kernelling', as mentioned earlier. In this way, complex data points can be separated with the help of more dimensions. The concept highlighted here is that the data points continue to get mapped into higher dimensions until a hyperplane is identified that shows a clear separation between the data points.

The figure below gives the 3D visualization of the above use case:

Types of Support Vector Machines

Support vector machines are broadly classified into two types: simple or linear SVM and kernel or non-linear SVM.

1. Simple or linear SVM

A linear SVM refers to the SVM type used for classifying linearly separable data. This implies that when a dataset can be segregated into categories or classes with the help of a single straight line, it is termed a linear SVM, and the data is referred to as linearly distinct or separable. Moreover, the classifier that classifies such data is termed a linear SVM classifier.

A simple SVM is typically used to address classification and [regression analysis](#) problems.

2. Kernel or non-linear SVM

Non-linear data that cannot be segregated into distinct categories with the help of a straight line is classified using a kernel or non-linear SVM. Here, the classifier is referred to as a non-linear classifier. The classification can be performed with a non-linear data type by adding features into higher dimensions rather than relying on 2D space. Here, the newly added features fit a hyperplane that helps easily separate classes or categories.

Kernel SVMs are typically used to handle optimization problems that have multiple variables.

Examples of Support Vector Machines

SVMs rely on supervised learning methods to classify unknown data into known categories. These find applications in diverse fields. Here, we'll look at some of the top real-world examples of SVMs:

1. Addressing the geo-sounding problem

The geo-sounding problem is one of the widespread use cases for SVMs, wherein the process is employed to track the planet's layered structure. This entails solving the inversion problems where the observations or results of the issues are used to factor in the variables or parameters that produced them.

In the process, linear function and support vector algorithmic models separate the electromagnetic data. Moreover, linear programming practices are employed while developing the supervised models in this case. As the problem size is considerably small, the dimension size is inevitably tiny, which accounts for mapping the planet's structure.

2. Assessing seismic liquefaction potential

Soil liquefaction is a significant concern when events such as earthquakes occur. Assessing its potential is crucial while designing any civil infrastructure. SVMs play a key role in determining the occurrence and non-occurrence of such liquefaction aspects.

Technically, SVMs handle two tests: SPT (Standard Penetration Test) and CPT (Cone Penetration Test), which use field data to adjudicate the seismic status.

Moreover, SVMs are used to develop models that involve multiple variables, such as soil factors and liquefaction parameters, to determine the ground surface strength. It is believed that SVMs achieve an accuracy of close to 96-97% for such applications.

3. Protein remote homology detection

Protein remote homology is a field of computational biology where proteins are categorized into structural and functional parameters depending on the sequence of amino acids when sequence identification is seemingly difficult. SVMs play a key role in remote homology, with kernel functions determining the commonalities between protein sequences.

Thus, SVMs play a defining role in computational biology.

4. Data classification

SVMs are known to solve complex mathematical problems. However, smooth SVMs are preferred for data classification purposes, wherein smoothing techniques that reduce the data outliers and make the pattern identifiable are used.

Thus, for optimization problems, smooth SVMs use algorithms such as the Newton-Armijo algorithm to handle larger datasets that conventional SVMs cannot. Smooth SVM types typically explore math properties such as strong convexity for more straightforward data classification, even with non-linear data.

5. Facial detection & expression classification

SVMs classify facial structures vs. non-facial ones. The training data uses two classes of face entity (denoted by +1) and non-face entity (denoted as -1) and $n \times n$ pixels to distinguish between face and non-face structures. Further, each pixel is analyzed, and the features from each one are extracted that denote face and non-face characters.

Finally, the process creates a square decision boundary around facial structures based on pixel intensity and classifies the resultant images. Moreover, SVMs are also used for facial expression classification, which includes expressions denoted as happy, sad, angry, surprised, and so on.

6. Surface texture classification

In the current scenario, SVMs are used for the classification of images of surfaces. Implying that the images clicked of surfaces can be fed into SVMs to determine the texture of surfaces in those images and classify them as smooth or gritty surfaces.

7. Text categorization & handwriting recognition

Text categorization refers to classifying data into predefined categories. For example, news articles contain politics, business, the stock market, or sports. Similarly, one can segregate emails into spam, non-spam, junk, and others.

Technically, each article or document is assigned a score, which is then compared to a predefined threshold value. The article is classified into its respective category depending on the evaluated score.

For handwriting recognition examples, the dataset containing passages that different individuals write is supplied to SVMs. Typically, SVM classifiers are trained with sample data initially and are later used to classify handwriting based on score values. Subsequently, SVMs are also used to segregate writings by humans and computers.

8. Speech recognition

In speech recognition examples, words from speeches are individually picked and separated. Further, for each word, certain features and characteristics are extracted. Feature extraction techniques include Mel Frequency Cepstral Coefficients (MFCC), Linear Prediction Coefficients (LPC), Linear Prediction Cepstral Coefficients (LPCC), and others.

These methods collect audio data, feed it to SVMs and then train the models for [speech recognition](#).

9. Stenography detection

With SVMs, you can determine whether any digital image is tampered with, contaminated, or pure. Such examples are helpful when handling security-related matters for organizations or government agencies, as it is easier to encrypt and embed data as a watermark in high-resolution images.

Such images contain more pixels; hence, it can be challenging to spot hidden or watermarked messages. However, one solution is to separate each pixel and store data in different datasets that SVMs can later analyze.

10. Cancer detection

Medical professionals, researchers, and scientists worldwide have been toiling hard to find a solution that can effectively detect cancer in its early stages. Today, several [AI](#) and ML tools are being deployed for the same. For example, in January 2020, Google developed an AI tool that helps in early breast cancer detection and reduces false positives and negatives.

In such examples, SVMs can be employed, wherein cancerous images can be supplied as input. SVM algorithms can analyze them, train the models, and eventually categorize the images that reveal malign or benign cancer features.

Takeaway

SVMs are crucial while developing applications that involve the implementation of predictive models. SVMs are easy to comprehend and deploy. They offer a sophisticated machine learning algorithm to process linear and non-linear data through kernels.

SVMs find applications in every domain and real-life scenarios where data is handled by adding higher dimensional spaces. This entails considering factors such as the tuning hyper-parameters, selecting the kernel for execution, and investing time and resources in the training phase, which help develop the supervised learning models.