Inventado originalmente por Vladimir Vapnik e Alexey Chervonenkis em 1963 com modificações para permitir classificações não-lineares em 1992 por Vapnik, os SVM são modelos de aprendizagem supervisionada utilizados nas tarefas de classificação e regressão.

Na SVM, os dados são representados por pontos num espaço n-dimensional, sendo n o número de atributos dos seus dados. SVM são classificadores binários lineares pois tenta buscar um hiperplano que separe linearmente os dados em dois grupos distintos. Podem existir diversos hiperplanos de separação, porém a melhor escolha é o plano que possui maior margem, ou seja, o plano que contém a maior separação entre os dados mais próximos, dessa forma diminuindo o erro de generalização do classificador. Novos elementos que forem classificados serão julgados de acordo com qual lado do hiperplano eles se encontram.

Casos que não podem ser linearmente separados podem ser modificados através das funções de kernel: funções que convertem os dados em espaços de maiores dimensões onde a análise pode ser feita e os dados separados linearmente por um hiperplano. Este hiperplano encontrado pode não ser uma reta nos planos originais dos dados.

São algoritmos que funcionam bem em espaços de muitas dimensões, até mesmo maior do que o número de exemplos da base de dados, porém não são eficientes quando os dados possuem muito ruído.

Em problemas de regressão as técnicas principais do SVM ainda podem ser usadas. Um hiperplano pode ser encontrado que divide o conjunto de dados em 2, porém a margem de separação passa a ser uma margem de erro aceitável, sendo que os dados que caiam dentro dessa margem de erro podem ser calculados seus valores a partir do hiperplano de separação e os dados que caem fora das margens são descartados como pontos fora da curva.