УДК 811.111:004.8

SUPPORT VECTOR MACHINE: THE MACHINE LEARNING ALGORITHM

Lozichniy A. S., student
Scientific supervisor – Vanik I.Y., senior lecturer
English language department №1
Belarusian National University of Technology
Minsk, Republic of Belarus

Support Vector Machines (SVM), introduced by Vladimir Vapnik and Alexey Chervonenkis in 1992, belong to a family of supervised learning algorithms primarily employed for classification and regression tasks. These algorithms aim to identify an ideal hyperplane dividing diverse groups in an n-dimensional space, thus enabling accurate predictions concerning unknown data points. Over recent years, SVMs have garnered considerable attention owing to their remarkable ability to manage high-dimensional data along with impressive efficiency across varying contexts. This article aims to delineate the core principles governing SVMs, elucidating their merits, demerits, extensions, competing strategies, and eventual implications.

One of the key strengths of SVM is its ability to handle non-linear relationships through the use of kernel functions. These functions transform the input data into a higher-dimensional space where a linear separation is possible. Common kernel functions include linear, polynomial, radial basis function (RBF), and sigmoid. SVM has several advantages over other machine learning algorithms. It is effective in high-dimensional spaces, even when the number of dimensions exceeds the number of samples. SVM is also memory efficient, as it only uses a subset of training points called support vectors to define the decision boundary.

At heart, SVM seeks to construct a model capable of discerning whether a provided sample stems from either Category A or Category B via discovering the most suitable border - referred to as a hyperplane - demarcating these groupings inside the p-dimensional real coordinate space (R^p). By situating this division equidistant from separate clusters, we optimize the margin – defined as the shortest distance amidst any point within each cluster and our chosen hyperplane. Such positioning enhances the likelihood of successful generalizations pertaining to previously undiscovered observations, ultimately minimizing overfitting concerns.

Furthermore, contrary to simple linear boundaries, employing kernel functions facilitates SVM models to tackle nonlinear challenges proficiently. Through mapping low-dimensional inputs onto elevated feature spaces, ostensibly complicated non-separable entities transform into linearly distinguishable ones [1]. In addition to classification tasks, SVM can also be used for regression by modifying the loss function to minimize deviations from a given target value rather than predicting class labels. SVM excel in their adaptability to diverse data types, robustness against overfitting, competence in high-dimensional spaces, proficiency in handling nonlinear relationships

However, SVM often struggle to complicate some ambiguous problem. For example, exorbitantly extensive repositories occasionally impede direct application of SVM due to inherent memory restrictions as well as memory capacity and restriction where archiving every instructional specimen might not consistently prove tenable, notably when confronted with monumental databanks.

Researchers have developed refined versions of classic SVM like LS-SVM, Nu-SVC, RAE, and WSVM to address computational challenges, optimization difficulties, customization needs, and performance enhancements in various scenarios. Also there are various alternatives to SVM that much powerful and can replace it one and for all. Most of them are already well-known in limited circles like Decision Trees (DT), Random Forests (RF) and Naive Bayes (NB) which offer alternative solutions to similar challenges with varying trade-offs in interpretability, precision, computational requirements, and assumptions.

Overall, SVM is a versatile and powerful machine learning algorithm that can be applied to a wide range of problems. Its ability to handle high-dimensional data, non-linear relationships, and its flexibility in parameter tuning make it a popular choice for many practitioners in the field of machine learning.

References

1. Machine Learning Algorithms – A Review [Electronic resource] – Mode of access: https://www.hindawi.com/journals/mpe/2021/5594899/. – Date of access: 10.03.2024.