NTNU
Norwegian University of Science and Technology

Regression with Support Vector Machines (SVM)

Week 03 – Advanced Topic 3

Agenda

- Introduction
- Objectives
- Theory
- Results
- Conclusions

NTNU

Introduction

- Type of Support Vector Machine (SVM).
- Supervised learning algorithm:
  - Classification problems.
- SVM identifies classes (hyperplane).
- Robust to outliers.
- Effective in high-dimensions.

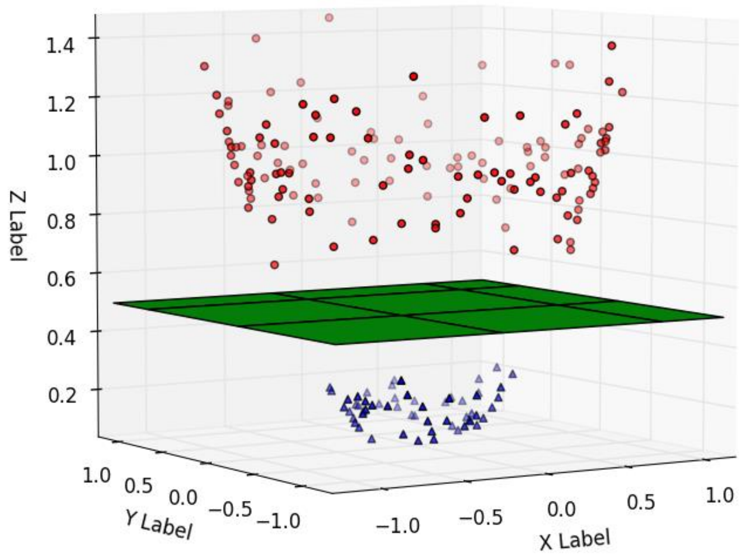


Fig.1 Hyperplane visualization (green) in 3D space.
(source:https://images.app.goo.gl/gifF9Bcss1o6vBJAA)

NTNU

Objectives

- SVM utilization for regression analysis.
- Estimate the functions f(x) such that:
  - Define the range of deviation.
- Optimal hyperplane identification:
  - Maximize margin (~ε-insensitive tube).
  - Distance between hyperplane and closest data (support vectors).
  - Tune hyper parameters (C, ε).
- Define error tolerance (ε-insensitive tube).
- Find the best fit while reducing complexity to avoid **overfitting** (flatness*).

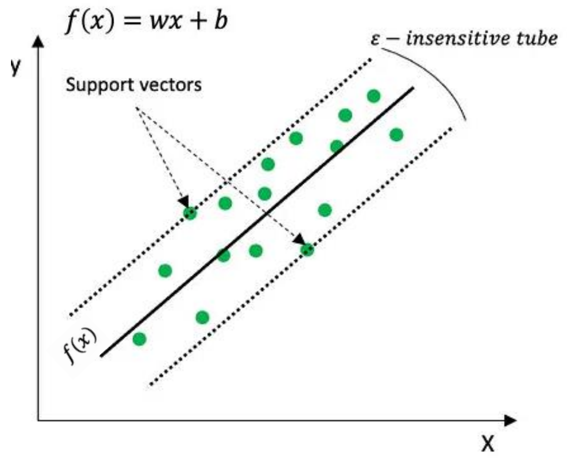
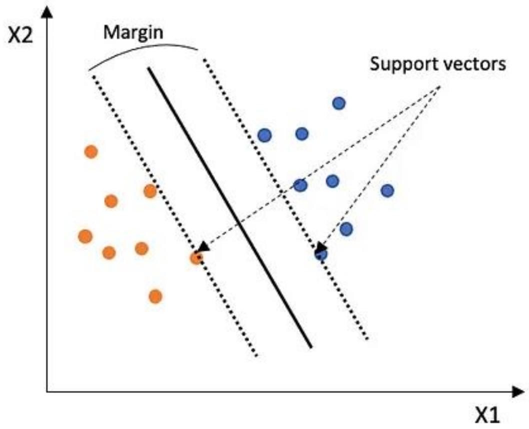
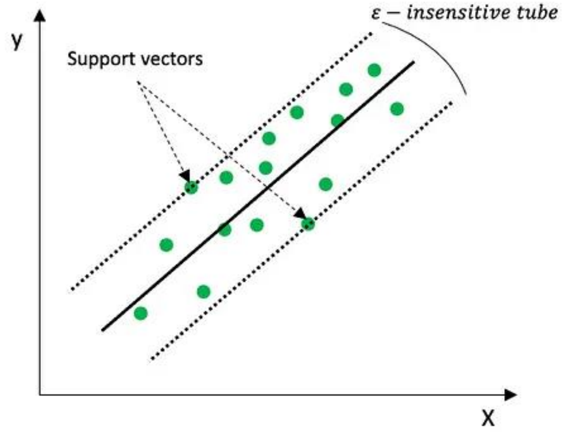


Fig.2 SVR hyperplane visualization in 2D space.

Theory



Classification problem using SVM



Regression problem using SVR

$$f(x) = \frac{1}{2}||w||^2 + C \sum_{\{i=1\}}^{\{n\}} (\xi_n + \xi_n^*)$$

Flatness of function

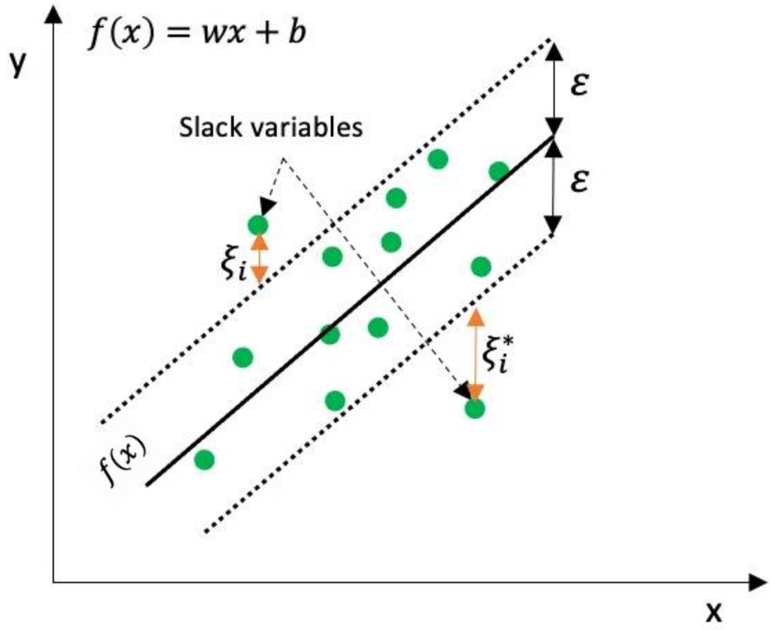
Regularization parameter
(Controls trade-off between low error and maintaining a flatness)

Penalty for errors
(Measures how far a point lies out of the ε-tube)

NTNU

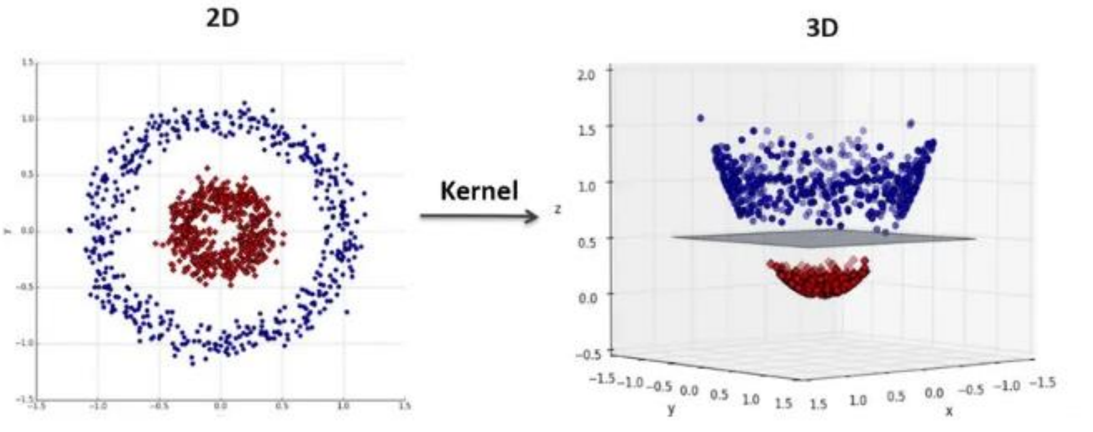
ε-insensitive tube

- ε is typically a hyperparameter.
- Permits some level of deviation.
- Data points within the tube are not penalized.
- Keeps balance between:
  - Model complexity (flatness*).
  - Generalization power.
- Outliers – Measured distance between data and support vector (ξi).



Theory

- Non-linear methods.
- Effective in high-dimensions.
- Robustness.
- Lower to higher dimension data transformation:
  - 2D to 3D transformation.
  - Data separation.
- Several types of kernels (e.g polynomial, RBF).



Complex in low dimensions

Simple in higher dimensions

Fig.3 Kernel transformation from lower to higher dimensions. (source:https://medium.com/@Suraj_Yadav/what-is-kernel-trick-in-svm-interview-questions-related-to-kernel-trick-97674401c48d)

NTNU

Results

- Univariate dataset (student marks).
  - Feature is the time_study.
- Goal is to predict:
  - Student marks given average time study/day.
- Different type of SVR were used :
  - "Linear"
  - "RBF"
  - "Polynomial"
- Hyper parameters:
  - C = 1.0
  - $\varepsilon = 0.1$

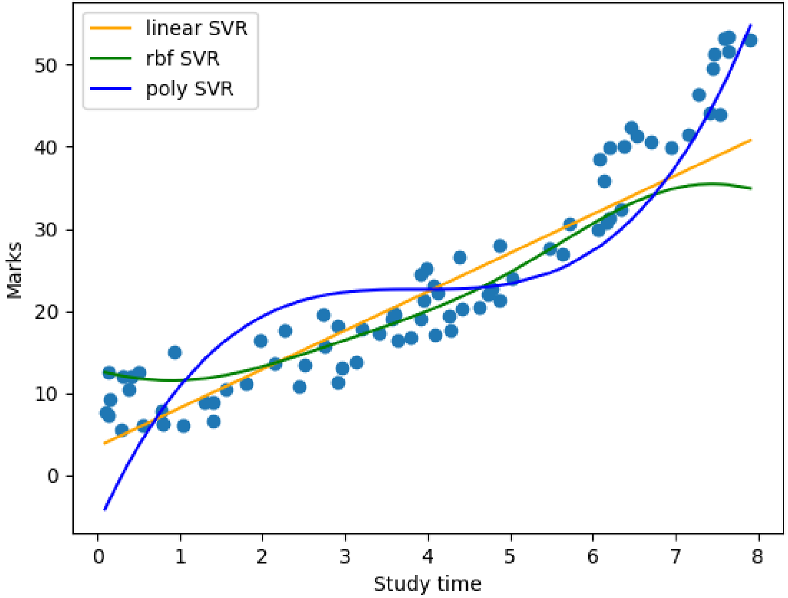


Fig.4 Fitted linear and non-linear SVR on train dataset

NTNU

Results

- Univariate dataset (student marks).
  - Feature is the time_study.
- Goal is to predict:
  - Student marks given average time study/day.
- Evaluation of "**Linear**" SVR:
  - R-squared (R2) coefficient: 0.83
  - RMSE : 6.7
- Evaluation of "**RBF**" SVR:
  - R-squared (R2) coefficient: 0.69
  - RMSE : 9.03
- Evaluation of "**Polynomial**" SVR:
  - R-squared (R2) coefficient:0.83
  - RMSE : 6.66

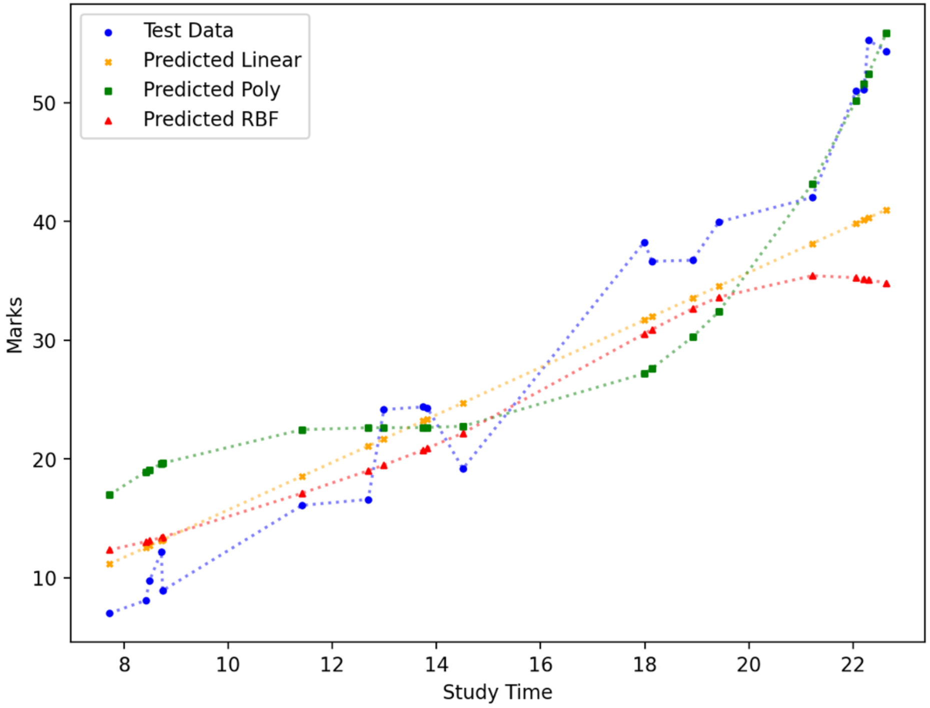


Fig.5 Test dataset and predictions from different SVRs

NTNU

Conclusions

- SVR extends the use of SVM to regression analysis.
- Defining ε-insensitive tube, increase some level of deviation.
- Tune parameters for optimal data "fit".
- Can handle linear and non-linear regression problems.
- Pros:
  - Can handle linear and non-linear data
  - Customizable – by tuning the parameter
  - Works well in high dimensional space.
- Cons:
  - No straight-forward way of choosing a kernel (only trial and error).
  - Complex kernels can overfit.
  - Large computational cost – takes time to train.
  - Difficult to explain when using non-linear kernels.

References

- https://www.analyticsvidhya.com/blog/2021/05/support-vector-machines/
- https://medium.com/@niousha.rf/support-vector-regressor-theory-and-coding-exercise-in-python-ca6a7dfda927
- *Patle, Arti, and Deepak Singh Chouhan. "SVM kernel functions for classification." 2013 International conference on advances in technology and engineering (ICATE). IEEE, 2013.*
- *Murty, M. N., et al. "Kernel-based SVM." Support vector machines and perceptrons: Learning, optimization, classification, and application to social networks (2016): 57-67.*
- https://www.kaggle.com/datasets/yasserh/student-marks-dataset
- https://roboticsbiz.com/pros-and-cons-of-support-vector-machine-svm/

NTNU

Thank you!

NTNU