MachineLearnAthon - Microlecture Forecasting Methods

Recorded by Lara Kuhlmann

Learning outcomes of today

After successfully completing this micro-lecture, you are able to….

- Name, describe and differentiate between methods for time-series forecasting
- Understand how to transform a time series as input for a machine learning model

Agenda for today

- (S)ARIMA
- Preparing time series for machine learning models
- Support Vector Machine
- Support Vector Regression
- Ensemble Model
- Gradient Boosting
- LightGBM

(S)ARIMA

- One of the most widely used approaches for time series forecasting
- **AR (AutoRegressive)**: Uses the relationship between an observation and a number of lagged observations (previous values).
- **I (Integrated)**: Involves differencing the raw observations to make the time series stationary (removing trends).
- **MA (Moving Average)**: Models the relationship between an observation and a residual error from a moving average model applied to lagged observations.
- SARIMA is an extension or ARIMA including seasonality

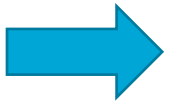
Hyndman, R.J., & Athanasopoulos, G. (2021) *Forecasting: principles and practice*, 3rd edition
A. E. Permanasari, I. Hidayah and I. A. Bustoni, "SARIMA (Seasonal ARIMA) implementation on time series to forecast the number of Malaria incidence" (*2013)*

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon

Machine Learning Methods for Time Series

- The standard time series method take only a time series as input
- For applying a machine learning algorithm, the time series needs to be transformed into a dataset with lags

t	y
2024-01-01	98
2024-02-01	103
2024-03-01	110
2024-04-01	87
2024-05-01	99
2024-06-01	93



Date	y	Lag_1	Lag_2	Lag_3
2024-01-01	98	NaN	NaN	NaN
2024-02-01	103	98	NaN	NaN
2024-03-01	110	103	98	NaN
2024-04-01	87	110	103	98
2024-05-01	99	87	110	103
2024-06-01	93	99	87	110

Support Vector Machine

- Solves binary classification problems
- Objective: find the maximum distanc between the hyperplane to classify th largest number of training points
- The data points of the hyperplane that ar closest to it are called “support vectors”

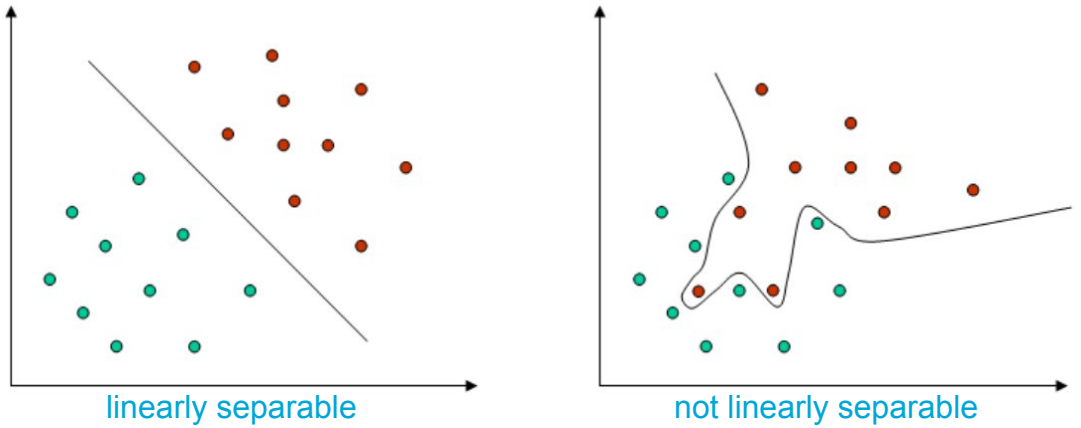


Image Source: https://de.wikipedia.org/wiki/Support_Vector_Machine

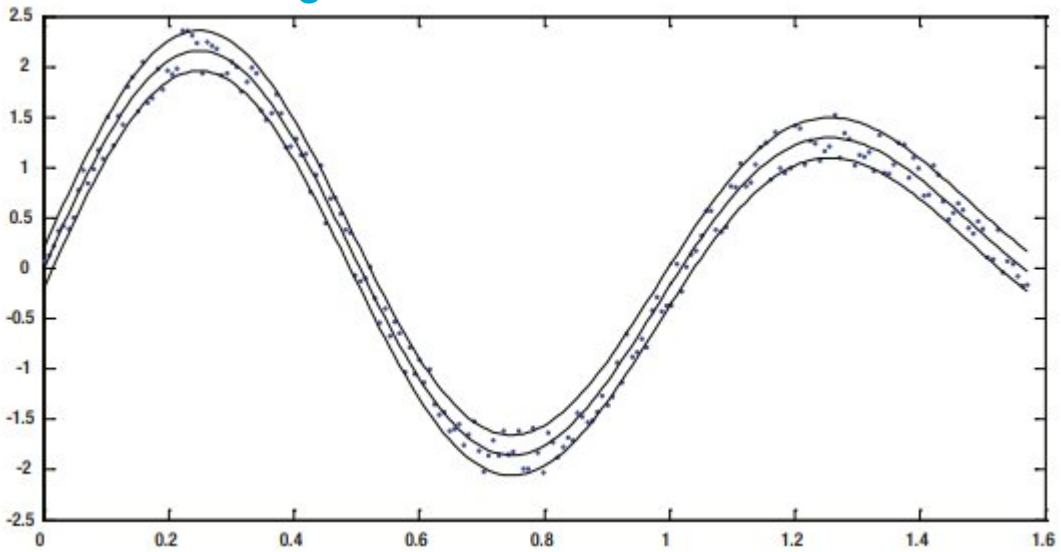
Mahdi Abolghasemi u.a. „Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion“. In: Compu ters & Industrial Engineering 142 (2020)
Mariette Awad und Rahul Khanna. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers (2015)

Co-funded by the Erasmus+ programme of the European Union

Support Vector Regression

- Support Vector Regression (SVR) is a generalization of the Support Vector Machine (SVM)
- Defined as an optimization problem and trained with a symmetric loss function
- During training, the loss function penalizes too high and too low misestimates
- The result of the optimization problem is the flattest tube, which contains the largest number of training instances

Nonlinear regression



Mariette Awad und Rahul Khanna. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers (2015);p. 73

Mahdi Abolghasemi u.a. „Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion“. In: Compu ters & Industrial Engineering 142 (2020)
Mariette Awad und Rahul Khanna. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers (2015)

Co-funded by the Erasmus+ programme of the European Union

Ensemble Model

- Combination of several models
- Improving prediction accuracy and reducing the risk of overfitting
- Bagging: ‘Bootstrap aggregation’
  - Uses different samples to train the decision trees
  - Result is the average of all predictions
  - Advantage: outliers have less influence
  - Example method: Random Forest
- Boosting:
  - Add iteratively weak learners to the model, which focus on data that has not yet been learned well
  - Higher risk of overfitting

Carsten Lanquillon. „Grundzüge des maschinellen Lernens“ (2019)
VKishore Ayyadevara. Pro Machine Learning Algorithms: A Hands-On Approach to Implementing Algorithms in Python and R. (2018)
Joos Korstanje. Advanced Forecasting with Python. With State-of-the Art-Models Including LSTMs, Facebook’s Prophet, and Amazon’s DeepAR (2021)

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon

Gradient Boosting

- Ensemble method, which uses boosting
- Suitable for regression and classification problems
- Sequential approach
  - Adds decision trees that try to reduce the error of the previous one
- Combination of weak learners into a strong model
- "Gradient": Error of the created model
- Goal: Generation of a small error
- Example: XGBoost and LightGBM

VKishore Ayyadevara. Pro Machine Learning Algorithms: A Hands-On Approach to Implementing Algorithms in Python and R (2018)
Sunitha Cheriyan u.a. „Intelligent Sales Prediction Using Machine Learning Techniques“ (2018)
Arnab Mitra u.a. „A Comparative Study for Machine Learning Models in Retail Demand Forecasting“(2023)

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon

LightGBM

- Light Gradient Boosting Machine
- Based on Gradient Boosting
- Uses GOSS and EFB
  - Gradient-Based One-Side Sample (GOSS): Filters out data points that have already been learnt
  - Exclusive Feature Bundling (EFB): Enables fast learning

Zharfan Akbar Andriawan u.a. „Vorhersage der Stornierung von Ho telbuchungen mit CRISP-DM“ (2020)
Joos Korstanje. Advanced Forecasting with Python. With State-of-the Art-Models Including LSTMs, Facebook’s Prophet, and Amazon’s DeepAR (2021)

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon

Recap this lecture

After successfully completing this lecture, you are able to….

- Name, describe and differentiate between methods for time-series forecasting
- Understand how to transform a time series as input for a machine learning model

The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

This material was created with the assistance of an AI language model.

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon