Machine Learning in Marketing–Theoretical Foundations and Applications

Syllabus WS 2020/21

Humboldt University Berlin, Institute of Marketing

Dr. Sebastian Gabel

Course Dates

Mondays, 17:00 – 20:00 (lecture and exercise), starting Nov. 9, 2020.

Course Prerequisites

No prerequisites, but successful participation in at least one of the following statistics/data science courses is recommended: Advanced Marketing Modeling, Selected Topics in Statistical and Machine Learning, Business Analytics & Data Science, Advanced Data Analytics for Management Support. In addition, students should be proficient in Python (or a comparable programming language) and have a good understanding of statistics and linear algebra. A basic knowledge of computer science is helpful. The course language is English.

Description and Objectives

This course is designed for master students from quantitative fields such as marketing, economics, statistics and computer science in their last year of study. It prepares students for solving real-world marketing problems using modern quantitative methods and is a good preparation for a machine learning/data science job in marketing or a PhD in quantitative marketing.

The course first reviews theoretical foundations in marketing, statistics, probability theory and computer science that are required to understand, apply and customize complex statistical models. The course will then focus on formalizing marketing decisions as machine learning problems and equips students with the necessary tools to efficiently implement machine learning models and pipelines. In the exercises, students will not only apply existing modelling libraries but also implement models and algorithms (almost) from scratch. After completing this course participants will be able to judge how modern machine learning methods complement (or even replace) traditional statistical methods for data analysis and decision-making.

The course content complements existing courses in that it reviews the theoretical foundations taught in statistics and computer science programs and then shows how to implement machine learning approaches to important marketing questions.

Grading

Grades are based on the final (written) exam. Successful participation in the home assignments serves as a preparation for the exam and is not mandatory for admission. Reference solutions to the home assignments will be discussed in the exercises.

Course Website

Moodle: https://moodle.hu-berlin.de/course/view.php?id=98455

Software

All lectures and exercises are based on Python. Students will use popular data and machine learning libraries, for example, numpy, scipy, pandas, scikit-learn, pytorch, lightgbm and statsmodels. The course participants are responsible for setting up and maintaining their development environment for following the lectures/exercising and solving the homework assignments.

Detailed Course Schedule (Preliminary)

CW	Date	Session 1	Session 2
45	02.11.2020	Dies Academicus	
46	09.11.2020	Org; Marketing Fundamentals	Introduction to Python and Colab
47	16.11.2020	Marketing Fundamentals (cntd.)	Homework 1 (“Python and Colab”)
48	23.11.2020	Machine Learning Fundamentals	Implementation 1 (configs, pipelines)
49	30.11.2020	Methodologies 1 (cluster analysis and dimensionality reduction)	Implementation 2 (sklearn)
50	07.12.2020	Methodologies 2 (logistic regression and tree-based methods)	Homework 2 (“ML Fundamentals & Implementation”)
51	14.12.2020	Application 1: Purchase Prediction with Boosted Trees	Implementation 3 (lightgbm); review homework task 3
Christmas break			
1	04.01.2021	Methodologies 3 (matrix factorization and neural matrix factorization)	Homework 3 (“Recommender Systems”)
2	11.01.2021	Methodologies 4 (neural networks)	Simulating Data (clustering)
3	18.01.2021	Guest Lecture 1 (academic speaker)	SupermarketGym; review homework task 4
4	25.01.2021	Application 2: Market Structure Analysis With Product Embeddings	Homework 4 (“Coupon Policies”)
5	01.02.2021	Application 3: Personalized Coupons With Deep Neural Networks	Implementation 4 (PyTorch); review homework task 5
6	08.02.2021	Application 4: Deep Recommender Systems	Guest Lecture 2 (industry speaker)
7	15.02.2021	Homework 5 (“P2V-MAP”)	
8	22.02.2021	Q&A	
8	26.02.2021	Exam	
CW = Calendar Week			

Readings

CW 48 (optional)	Bergstra, J. and Bengio, Y., 2012. Random search for hyper-parameter optimization. *The Journal of Machine Learning Research*, 13(1), pp.281-305.
CW 2 (optional)	LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep Learning. *Nature*, 521(7553), pp.436-444.
CW 3 (optional)	Neslin, S.A. and van Heerde, H.J., 2009. Promotion Dynamics. *Foundation and Trends in Marketing*, 3 (4), 177–268.
CW 4	Gabel, S., Guhl, D. and Klapper, D., 2019. P2V-MAP: Mapping Market Structures for Large Retail Assortments. *Journal of Marketing Research*, 56(4), pp.557-580.
CW 5	Gabel, S. and Timoshenko, A., 2020. Product Choice with Large Assortments: A Scalable Deep-Learning Model. Available at *SSRN*: https://ssrn.com/abstract=3402471.
CW 6	Covington, P., Adams, J. and Sargin, E., 2016, September. Deep Neural Networks for Youtube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 191-198).

Literature

Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Springer.

Friedman, J., Hastie, T. and Tibshirani, R., 2001. T*he Elements of Statistical Learning* (Vol. 1, No. 10). New York: Springer Series in Statistics.

Goodfellow, I., Bengio, Y. and Courville, A., 2016. *Deep Learning*. MIT press.

Murphy, K.P., 2012. *Machine Learning: A Probabilistic Perspective*. MIT press.

Palmatier, R.W. and Sridhar, S., 2017. *Marketing strategy: Based on first principles and data analytics*. Macmillan International Higher Education.

(See Moodle System for additional research articles.)