Список литераторы по курсу

«МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ»

2021

Лекция 1. Введение

Лекция 2. Теория вероятностей и статистика

* Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg. URL: <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>
* OpenIntro Statistics URL: <https://www.openintro.org/book/os/>
* Кремер Н.Ш. Глава 9. Основы математической теории выборочного метода // Теория вероятностей и математическая статистика

Лекция 3. Линейная регрессии и метод наименьших квадратов

* Chapter 3. Linear Regression// An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. URL: <https://www.statlearning.com/>

Лекция 4. Градиентный спуск и стохастический градиентный спуск

* Nocedal, Jorge and Wright, Stephen J.. Numerical optimization. 2. ed. New York, NY: Springer, 2006.
* Chapter 5. Logistic Regression // Speech and Language Processing by Daniel Jurafsky & James H. Martin. URL: <https://web.stanford.edu/~jurafsky/slp3/5.pdf>
* см. лекцию по распределенным алгоритмам

Лекция 5. Логистическая регрессия

* Chapter 5. Logistic Regression // Speech and Language Processing by Daniel Jurafsky & James H. Martin. URL: <https://web.stanford.edu/~jurafsky/slp3/5.pdf>
* Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg. URL: <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>

Лекция 6. Регуляризация

* Chapter 6. Linear Model Selection and Regularization // An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. URL: <https://www.statlearning.com/>
* Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg. URL: <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>

Лекция 7. Повторные выборки. Выбор модели

1. Chapter 5. Resampling Methods // An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. pp. 175–190. URL: <https://www.statlearning.com/>
2. Chapter 7. Model Assessment and Selection // The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani, Jerome Friedman. pp. 219–257. URL: <https://web.stanford.edu/~hastie/ElemStatLearn/>

Лекция 8. Выбор признаков

* Chapter 5. Resampling Methods // An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. pp. 203–264. URL: <https://www.statlearning.com/>

Лекция 9. Метод опорных векторов

* Zaki, Mohammed & Meira Jr, Wagner. Data Mining and Analysis: Fundamental Concepts and Algorithms. 2014.

Лекция 10. Наивный байесовский классификатор

* Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg. URL: <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>
* Text Classification using Naive Bayes by Hiroshi Shimodaira. URL: <https://www.inf.ed.ac.uk/teaching/courses/inf2b/learnnotes/inf2b-learn07-notes-nup.pdf>
* Chapter 13 Text classification and Naive Bayes // Introduction to Information Retrieval by Christopher D. Manning, Prabhakar Raghavan and Hinrich Schutze, Cambridge University Press. 2008. URL: <http://www-nlp.stanford.edu/IR-book/>

Лекция 10\*. Байесовская линейная регрессия

* Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg. URL: <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>

Лекция 11. Деревья решений

* Chapter 8. Tree-Based Methods // An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. URL: <https://www.statlearning.com/>
* Machine Learning FAQ. Why are we growing decision trees via entropy instead of the classification error? URL: <https://sebastianraschka.com/faq/docs/decisiontree-error-vs-entropy.html>
* Decision Tree (Data Mining). URL: <http://mlwiki.org/index.php/Decision_Tree_(Data_Mining)>

Лекция 12. Кластеризация

* Chapter 10.3 Clustering Methods // An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. URL: <https://www.statlearning.com/>
* Chapter 14.3. Cluster Analysis // The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani, Jerome Friedman. pp. 501–527. URL: <https://web.stanford.edu/~hastie/ElemStatLearn/>

Лекция 13. Метод главных компонент

1. Chapter 6. Dimension Reduction Methods // An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. pp. 228–237. URL: <https://www.statlearning.com/>
2. Chapter 10. Principal Components Analysis // An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshir. pp. 374–284. URL: <https://www.statlearning.com/>
3. Shlens, J. (2003). A tutorial on principal component analysis, URL: <https://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf>

Лекция 14. Рекомендательные системы

1. Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor. 2010. Recommender Systems Handbook (1st. ed.). Springer-Verlag, Berlin, Heidelberg
2. Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th international conference on World Wide Web (WWW ’01). Association for Computing Machinery, New York, NY, USA, 285–295. DOI:https://doi.org/10.1145/371920.372071
3. István Pilászy, Dávid Zibriczky, and Domonkos Tikk. Fast als-based matrix factorization for explicit and implicit feedback datasets. In Proceedings of the fourth ACM conference on Recommender systems (RecSys ’10). Association for Computing Machinery, New York, NY, USA, 71–78. DOI:https://doi.org/10.1145/1864708.1864726

Лекция 15. Распределённые алгоритмы

1. Zadeh, Reza & Carlsson, Gunnar. (2013). Dimension Independent Matrix Square using MapReduce.
2. CME 323: Distributed Algorithms and Optimization, Spring 2015 <http://stanford.edu/~rezab/dao>. Instructor: Reza Zadeh, Databricks and Stanford. Lecture 10, 4/29/2015. Scribed by Jan Dlabal, Fang-Chieh Chou, Yu-Wei Lin, Yi-Hong Kuo.
3. Lars Schmidt-Thieme. Lecture Note. Big Data Analytics. Distributed Machine Learning Algorithms: Distributed Stochastic Gradient Descent, University of Hildesheim, Germany
4. Niu, Feng & Recht, Benjamin & Ré, Christopher & Wright, Stephen. (2011). HOGWILD!: A Lock-Free Approach to Parallelizing Stochastic Gradient
5. R. Gemulla, P. J. Haas, E. Nijkamp, Y. Sismanis «Large-Scale Matrix Factorization with Distributed Stochastic Gradient Descent», IBM Research Report RJ10481, March 2011 Revised February, 2013
6. Yunhong Zhou, Dennis Wilkins, Robert Schreiber, Rong Pan «Large-Scale Parallel Collaborative Filtering for the Netflix Prize» AAIM '08 Proceedings of the 4th international conference on Algorithmic Aspects in Information and Management. PP 337–348, Springer-Verlag Berlin, Heidelberg, 2008