

Вопросы к государственному экзамену

Магистерская программа

«Логические и комбинаторные методы анализа данных»

1. Методы градиентного спуска и Ньютона для безусловной гладкой оптимизации. Стратегии выбора длины шага. Скорости сходимости методов.
2. Условия Каруша-Куна-Таккера для задач условной оптимизации. Двойственная задача оптимизации. Методы решения выпуклых условных задач с помощью метода Ньютона и метода логарифмических барьеров.
3. Понятие субградиента и субдифференциала, субдифференциальное исчисление. Субградиентный метод для решения негладких задач выпуклой оптимизации. Его скорость сходимости.
4. Проксимальный градиентный метод для минимизации композитных функций. Выбор длины шага. Скорость сходимости метода. Примеры применения.
5. Стохастический градиентный спуск, его скорость сходимости. Методы стохастической оптимизации с линейной скоростью сходимости: SAG и SVRG.
6. Функции ошибки в задаче регрессии (MAE, RMSE, SMAPE, MAPE, PMAD, MRAE, REL_MAE, PB, MASE).
7. Функционалы качества в задачах чёткой бинарной классификации и классификации на k (пересекающихся) классов (TPR, Specificity, Precision, F1, Cohen's Kappa, MCC, BA).
8. Скоринговые функции ошибки (Log Loss, MSE, Misclassification Loss, Exploss) и кривые (ROC, PR, GINI, Gain Curve, Lift Curve, Kolomogorov Smirnov chart) в машинном обучении.
9. Анализ социальных сетей: безмасштабные сети, модель малого мира, коэффициент кластеризации, важность-центральность вершины, PageRank.
10. Анализ социальных сетей: определение вероятности появления ребра.
11. Анализ социальных сетей: нахождение сообществ в графе.
12. Случайные леса и бустинг над деревьями (описание алгоритмов, современные реализации, настройка параметров, критерии расщепления)
13. Лемма Бёрнсайда. Цикловой индекс действия группы на множестве. Теорема Пойа (без доказательства). Применение циклового индекса и теоремы Пойа при решении комбинаторных задач.
14. Структурный подход в распознавании образов: составляющие и особенности подхода. Левенштейновская аппроксимация произвольного слова словом из регулярного языка. Лемма о порядке редакторских операций.
15. Числа Стирлинга II рода: определение, выражение через суммирование биномиальных коэффициентов (с выводом), вывод рекуррентной формулы. Числа Белла.
16. Модель Эрдёша-Реньи случайных графов. Наблюдения Барабаша-Альберт. Модель интернета Боллобаша-Риордана.
17. Векторные представления слов. Модели Skip-gram и CBOW. Их модификации, основанные на иерархическом мягком максимуме и негативном семплировании.
18. Задача разметки последовательности (sequence labeling). Модель linear-CRF. Обучение модели и нахождение оптимальной последовательности меток. Совместное использование linear-CRF и нейросетевых архитектур.
19. Задача переноса обучения (transfer learning). Основные модели переноса обучения для текстовых данных: ELMO, ULMFIT и BERT.
20. Алгоритмическая рандомизация, задача вычисления порядковых статистик и медианы.
21. Алгоритм быстрого преобразования Фурье, приложение к цифровой обработке сигналов.
22. $O(n \log n)$ алгоритм вычисления пересечений конечного множества отрезков на плоскости.
23. Динамическое программирование, алгоритм выделения наибольшей общей подпоследовательности для двух последовательностей.

24. Триангуляция Делоне и ее свойства. Алгоритм построения триангуляции Делоне.
25. Свёрточные нейронные сети. Основные используемые архитектуры. Решение задач семантической сегментации изображений и детекции объектов на изображениях.
26. Рекуррентные нейронные сети, основные архитектуры. Проблема затухающих и взрывающихся градиентов, методы борьбы с ней. Механизм внимания и его использование при решении задачи машинного перевода.
27. EM-алгоритм для обучения вероятностных моделей со скрытыми переменными. Модель вариационного автокодировщика. Трюк репараметризации.
28. Обучение с подкреплением. Примеры практических задач. Q-обучение. Модель DQN.
29. Подход Policy gradient. Алгоритмы Reinforce и A2C.
30. Обучение с подкреплением с добавлением энтропии. Метод Soft Actor-Critic.

Литература

1. J. Nocedal, S.J. Wright. Numerical Optimization. Springer, 2006.
2. S. Boyd, L. Vandenberghe. Convex Optimization, Cambridge University Press, 2004.
3. D. Bertsekas. Convex Analysis and Optimization, Athena Scientific, 2003.
4. Optimization for Machine Learning. Edited by Suvrit Sra, Sebastian Nowozin and Stephen J. Wright, MIT Press, 2011.
5. M. Schmidt. Notes on Big-n Problems, 2012.
6. R. Johnson, T. Zhang. Accelerating Stochastic Gradient Descent using Predictive Variance Reduction // NIPS, 2013.
7. M. Schmidt, N. Le Roux, F. Bach. Minimizing Finite Sums with the Stochastic Average Gradient // ArXiv:1309.2388, 2013.
8. Стрижов В.В. Функция ошибки в задачах восстановления регрессии // Заводская лаборатория, 2013, 79(5): 65-73.
9. К.Д. Маннинг, П. Рагхаван, Х. Шютце «Введение в информационный поиск» // . — Вильямс, 2011.
10. D. Easley and J. Kleinberg. "Networks, Crowds, and Markets: Reasoning About a Highly Connected World".
11. S. Fortunato. Community detection in graphs. Physics Reports, Vol. 486, pp. 75-174, 2010
12. И. Генрихов О критериях ветвления, используемых при синтезе решающих деревьев // Машинное обучение и анализ данных, 2014, Т.1, №8, С.988-1017
13. A. Natekin, A. Knoll Gradient boosting machines, a tutorial // Front Neurorobot. 2013; 7: 21.
14. Белов В. В., Воробьев Е. М., Шаталов В. Е. Теория графов. М.: Высшая школа, 1976.
15. Гуров С. И. Конспект лекций по курсу «Комбинаторные и логические методы анализа данных». Электронный ресурс [http://www.machinelearning.ru/wiki/index.php?title=Комбинаторные_и_логические_методы_анализа_данных_%28курс_лекций%2C_С.И._Гуров%29
16. Райгородский А. М. Модели Интернета: учебное пособие. Долгопрудный: Издательский дом «Интеллект», 2013.
17. D. Jurafsky and J. H. Martin, Speech and Language Processing (3rd ed. draft), <https://web.stanford.edu/~jurafsky/slp3/>
18. T. Mikolov, K. Chen, G. Corrado, J. Dean, Distributed Representations of Words and Phrases and their Compositionality, <https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>
19. C. Sutton, A. McCallum, An introduction to Conditional Random Fields (2011), <https://homepages.inf.ed.ac.uk/csutton/publications/crftut-fnt.pdf>
20. G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, C. Dyer, Neural Architectures for Named Entity Recognition, <https://arxiv.org/abs/1603.01360>
21. M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, L. Zettlemoyer, Deep contextualized word representations, <https://arxiv.org/pdf/1802.05365.pdf>
22. J. Howard, S. Ruder, Universal Language Model Fine-tuning for Text Classification, <https://arxiv.org/abs/1801.06146>
23. J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, <https://arxiv.org/abs/1810.04805>
24. T. Hofmann, Probabilistic latent semantic indexing, <http://dataera.org/2014/04/notes-on-em-and-plsa/>
25. Кормен Т.Х. и др. Алгоритмы: построение и анализ, 3-е изд., Москва, «И. Д. Вильямс», 2016. — 1328 с.

26. Дасгупта С., Пападимитриу Х., Вазирани У. Алгоритмы. Москва, МЦНМО, 2014.
27. Препарата Ф., Шеймос М. Вычислительная геометрия: введение, Москва, «Мир». 1989. – 478 с.
28. Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning, MIT Press, 2016.
29. Abhishek Chaurasia, Eugenio Culurciello LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation, 2017. <https://arxiv.org/abs/1707.03718>
30. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // ArXiv: 1506.01497, 2015.
31. Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio. Neural Machine Translation by Jointly Learning to Align and Translate // ArXiv: 1409.0473, 2014.
32. Diederik P Kingma, Max Welling Auto-Encoding Variational Bayes, 2013. <https://arxiv.org/abs/1312.6114>
33. R. Sutton, A. Barto. Reinforcement Learning: An Introduction, 2nd ed., 2018.
34. Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, Martin Riedmiller. Playing Atari with Deep Reinforcement Learning, 2013. <https://arxiv.org/abs/1312.5602>
35. Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy P. Lillicrap, Tim Harley, David Silver, Koray Kavukcuoglu. Asynchronous Methods for Deep Reinforcement Learning, 2016. <https://arxiv.org/abs/1602.01783>

36. Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, Daan Wierstra. Continuous control with deep reinforcement learning, 2015. <https://arxiv.org/abs/1509.02971>
37. Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, Sergey Levine Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor, 2018. <https://arxiv.org/abs/1801.01290>