

Список вопросов к экзамену по курсу «Глубинное обучение», осень 2017

1. Стохастическая оптимизация. Стохастический градиентный спуск, метод Adagrad, метод ADAM. [1, глава 8; 2; 3]
2. Автоматическое дифференцирование: проход вперёд и назад. Вычисление произведения гессиана на произвольный вектор. Алгоритм обратного распространения ошибки. [4, глава 8.2; 5]
3. Сети прямого распространения. Модель автокодировщика. Примеры применения. Регуляризация в глубоких сетях: Dropout, Batch Normalization. [1, главы 6,7; 10]
4. Свёрточные нейронные сети. Модели AlexNet, VGG, Inception, ResNet. [1, глава 9; 11]
5. Локализация и детекция объектов на изображении. Методы R-CNN, Fast R-CNN, Faster R-CNN. [6]
6. Рекуррентные нейронные сети, процедура обучения. Проблема затухающих и взрывающихся градиентов, способы её решения. Модели LSTM, GRU. Применение рекуррентных сетей для решения практических задач. [1, глава 10]
7. Решение задачи машинного перевода. Модель Seq2seq. Механизм внимания. [12, 13]
8. Вероятностные модели со скрытыми переменными, EM-алгоритм. Вероятностная модель главных компонент. [14, глава 9; 15]
9. Модель вариационного автокодировщика. Трюк репараметризации [8]
10. Перенесение стиля на изображениях [7]
11. Обучение с подкреплением. Примеры практических задач. Q-обучение. Модель DQN. [9]
12. Обучение политики в обучении с подкреплением. Алгоритм REINFORCE. Подход Actor-Critic.
13. Генеративно-согласительные сети. Модель DCGAN. Примеры применения. [16, 17]
14. Задача структурного предсказания. Объединение структурного метода опорных векторов и нейронных сетей для задачи классификации последовательностей. [20]

Литература:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning, MIT Press, 2016.
2. J. Duchi, E.Hazan, Y.Singer. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization // JMLR, 2011
3. D. Kingma, J. Ba. Adam: A Method for stochastic optimization // ArXiv 1412.6980.
4. J. Nocedal, S.Wright. Numerical optimization. Springer, 2006.
5. N. Schraudolph. Fast Curvature Matrix-Vector Products for Second-Order Gradient Descent // Neural Computation, V.14, 2002, pp. 1723–1738.
6. S. Ren, K. He, R. Girshick, J. Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // ArXiv 1506.01497
7. L. Gatys, A. Ecker, M. Bethge. A Neural Algorithm of Artistic Style // ArXiv 1508.06576
8. D. Kingma, M. Welling. Autoencoding variational Bayes // ArXiv 1312.6114
9. V. Mnih et al. Playing Atari with Deep Reinforcement Learning.
10. S. Ioffe, C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariance shift // ArXiv 1502.03167
11. K. He, X. Zhang, S. Ren, J. Sun. Deep residual learning for image recognition // ArXiv 1512.03385
12. D. Bahdanau, K. Cho, Y. Bengio. Neural Machine Translation by Jointly Learning to Align and Translate // ArXiv 1409.0473
13. I. Sutskever, O. Vinyals, Q. Le. Sequence to Sequence Learning with Neural Networks // ArXiv 1409.3215
14. Bishop C. Pattern Recognition and Machine Learning, 2006.
15. Tipping M., Bishop C. Probabilistic Principal Component Analysis // Journal of the Royal Statistical Society B, 1999.
16. I. Goodfellow et al. Generative Adversarial Nets // ArXiv 1406.2661, 2014.

17. A. Redford et al. Unsupervised representation learning with deep convolutional generative adversarial networks // ICLR 2016.
18. Jang et al. "Categorical reparameterization with Gumbel-Softmax", 2017
19. Maddison et al. "The concrete distribution: A continuous relaxation of discrete random variables", 2017
20. M. Jaderberg, K. Simonyan, A. Vedaldi, A. Zisserman. Deep structured output learning for unconstrained text recognition // ArXiv 1412.5903, 2014.

Теоретический минимум

Вопросы из этой части охватывают базовые математические понятия и алгоритмы, которые активно используются в курсе. Незнание ответа на любой вопрос из данной части автоматически влечёт за собой неудовлетворительную оценку по экзамену.

1. Стохастический градиентный спуск.
2. Алгоритм обратного распространения ошибки.
3. Стандартные модели глубинного обучения: полносвязная сеть, свёрточная сеть, LSTM
4. Метод регуляризации dropout, батч нормализация.
5. EM-алгоритм для обучения вероятностных моделей со скрытыми переменными.
6. Трюк производной логарифма и трюк репараметризации.
7. Трюк Gumbel-softmax.
8. Алгоритм Q-обучения.
9. Алгоритм Reinforce.
10. Генеративно-сопоставительная сеть.