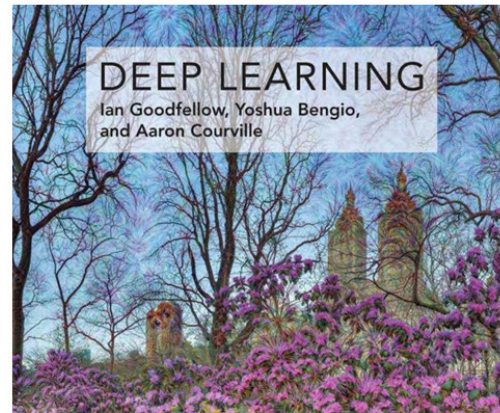
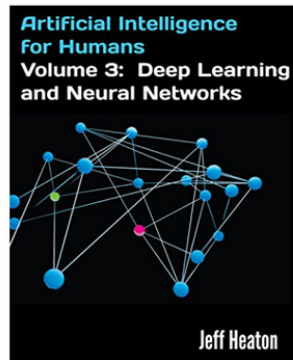
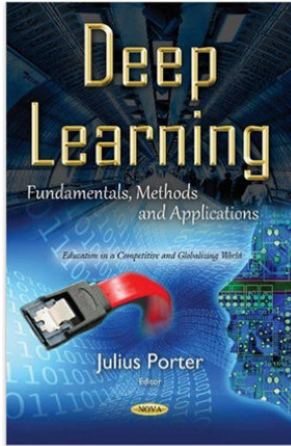









Apostila Teórica de Deep Learning

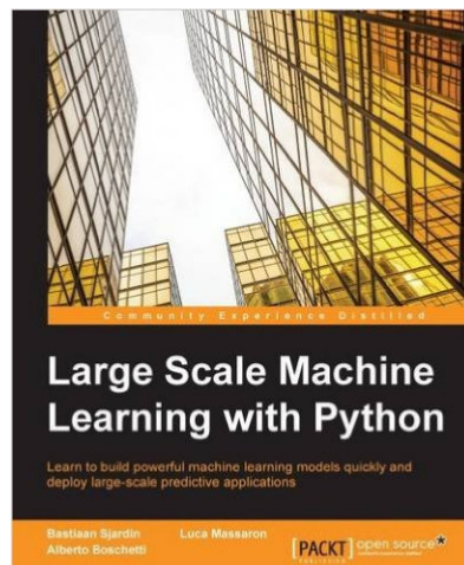
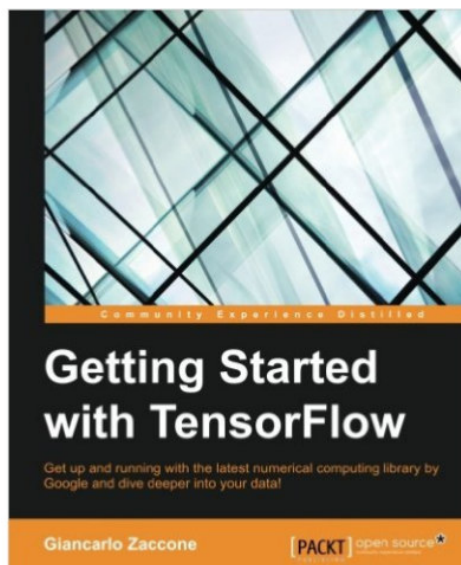
Transparência 01: Referências Bibliográficas Teóricas (1/2)



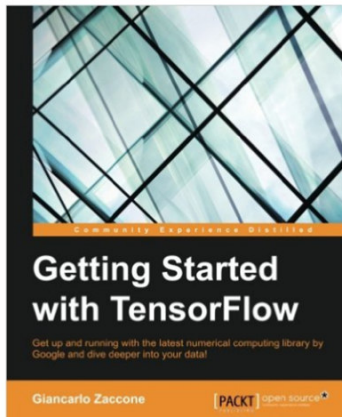
Transparência 02: Referências Bibliográficas Teóricas (2/2)

-  A Fast Learning Algorithm for Deep Belief Nets (PAPER), Hinton et al, 2006 .pdf
-  A Fast Learning Algorithm for Deep Belief Nets, Hinton et al, 2006 .pdf
-  Big Data Deep Learning Challenges and Perspectives, Chen and Lin, 2014.pdf
-  Deep Learning in Neural Networks An Overview, Schmidhuber J, 2015.pdf
-  Deep Learning Using Genetic Algorithms, Lamos-Sweeney, 2012.pdf
-  Deep Learning Using Linear Support Vector Machines, Tang, 2013.pdf
-  Deep Support Vector Machines for Regression Problems, Wiering et al, 2013.pdf

Transparência 03: Referências Bibliográficas Práticas (1/4)

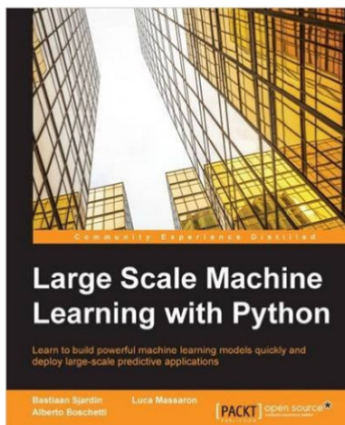


Transparência 04: Referências Bibliográficas Práticas (2/4)



Chapter 1, TensorFlow – Basic Concepts, contains general
Chapter 2, Doing Math with TensorFlow, describes the ability of mathematical processing of TensorFlow.
Chapter 3, Starting with Machine Learning, introduces some
Chapter 4, Introducing Neural Networks, provides a quick and detailed introduction of neural networks.
Chapter 5, Deep Learning, gives an overview of deep learning algorithms. Only in recent years has deep learning collected a large number of results considered unthinkable a few years
Chapter 6, GPU Programming and Serving with TensorFlow.

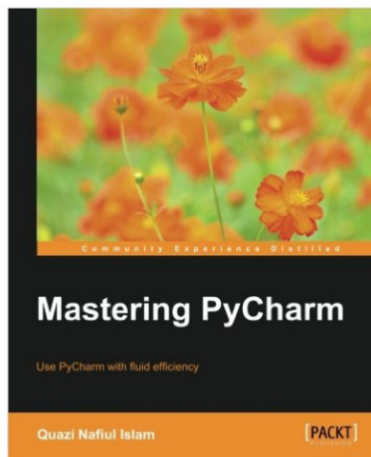
Transparência 05: Referências Bibliográficas Práticas (3/4)



Chapter 3, *Fast-Learning SVMs*, covers streaming algorithms that are capable of discovering non-linearity in the form of support vector machines.

Chapter 4, *Neural Networks and Deep Learning*, provides useful tactics for applying deep neural networks within the Theano framework together with large-scale applications with **Chapter 5**, *Deep Learning with TensorFlow*, covers interesting deep learning techniques together with an online method for neural networks. Although TensorFlow is only in its infancy, the framework provides elegant machine learning solutions.

Transparência 06: Referências Bibliográficas Práticas (4/4)



Transparência 07: Conteúdo do Curso de *Deep Learnig*

Confecção de quatro Apostilas, escritas em língua portuguesa, contendo os seguintes tópicos:

1. Introdução à Linguagem Python;
2. Introdução à IDE para Python conhecida como PyCharm;
3. Introdução à Plataforma TensorFlow;
4. Desenvolvimento teórico das redes neurais profundas (*Deep Neural Network*).

Transparência 08: Artigo Histórico Bastante Importante

Deep Learning in Neural Networks: An Overview

Technical Report IDSIA-03-14 / arXiv:1404.7828 v4 [cs.NE] (88 pages, 888 references)

Jürgen Schmidhuber
The Swiss AI Lab IDSIA
Istituto Dalle Molle di Studi sull'Intelligenza Artificiale
University of Lugano & SUPSI
Galleria 2, 6928 Manno-Lugano
Switzerland

8 October 2014

Transparência 09: Principais Tipos de Redes Neurais Profundas (*Deep Neural Networks*)

Deep Learning está ganhando mais e mais popularidade devido ao seu sucesso alcançado em várias aplicações como em Processamento de Linguagem Natural (*Natural Language Processing*, NLP), Reconhecimento de Imagens e outros paradigmas de Aprendizado de Máquinas (*Machine Learning*, ML).

Existem Três abordagens convencionais que formam as bases para a compreensão teórica adequada de *Deep Learning*:

1. *Convolutional Neural Networks (CNNs)* proposta por Lecun;
2. *Deep Belief Networks (DBNs)* proposta por Hinton;
3. *Stacked Auto-encoders* proposta por Bengio.

Transparência 10: *Convolutional Neural Networks (CNNs)* proposta por Lecun

- LeCun, Y. (1985). Une procédure d'apprentissage pour réseau à seuil asymétrique. *Proceedings of Cognitiva 85, Paris*, pages 599–604.
- LeCun, Y. (1988). A theoretical framework for back-propagation. In Touretzky, D., Hinton, G., and Sejnowski, T., editors, *Proceedings of the 1988 Connectionist Models Summer School*, pages 21–28, CMU, Pittsburgh, Pa. Morgan Kaufmann.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Back-propagation applied to handwritten zip code recognition. *Neural Computation*, 1(4):541–551.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1990a). Handwritten digit recognition with a back-propagation network. In Touretzky, D. S., editor, *Advances in Neural Information Processing Systems 2*, pages 396–404. Morgan Kaufmann.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- LeCun, Y., Denker, J. S., and Solla, S. A. (1990b). Optimal brain damage. In Touretzky, D. S., editor, *Advances in Neural Information Processing Systems 2*, pages 598–605. Morgan Kaufmann.
- LeCun, Y., Muller, U., Cosatto, E., and Flepp, B. (2006). Off-road obstacle avoidance through end-to-end learning. In *Advances in Neural Information Processing Systems (NIPS 2005)*.
- LeCun, Y., Simard, P., and Pearlmutter, B. (1993). Automatic learning rate maximization by on-line estimation of the Hessian's eigenvectors. In Hanson, S., Cowan, J., and Giles, L., editors, *Advances in Neural Information Processing Systems (NIPS 1992)*, volume 5. Morgan Kaufmann Publishers, San Mateo, CA.

Transparência 11: *Deep Belief Networks (DBNs)* proposta por Hinton

- Hinton, G. and Salakhutdinov, R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507.
- Hinton, G. E. (1989). Connectionist learning procedures. *Artificial intelligence*, 40(1):185–234.
- Hinton, G. E. (2002). Training products of experts by minimizing contrastive divergence. *Neural Comp.*, 14(8):1771–1800.
- Hinton, G. E., Dayan, P., Frey, B. J., and Neal, R. M. (1995). The wake-sleep algorithm for unsupervised neural networks. *Science*, 268:1158–1160.
- Hinton, G. E., Deng, L., Yu, D., Dahl, G. E., Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., and Kingsbury, B. (2012a). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Mag.*, 29(6):82–97.
- Hinton, G. E. and Ghahramani, Z. (1997). Generative models for discovering sparse distributed representations. *Philosophical Transactions of the Royal Society B*, 352:1177–1190.
- Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554.
- Hinton, G. E. and Sejnowski, T. E. (1986). Learning and relearning in Boltzmann machines. In *Parallel Distributed Processing*, volume 1, pages 282–317. MIT Press.
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2012b). Improving neural networks by preventing co-adaptation of feature detectors. Technical Report arXiv:1207.0580.
- Hinton, G. E. and van Camp, D. (1993). Keeping neural networks simple. In *Proceedings of the International Conference on Artificial Neural Networks, Amsterdam*, pages 11–18. Springer.

Transparência 12: *Stacked Auto-encoders* proposta por Bengio

- Bengio, Y. (1991). *Artificial Neural Networks and their Application to Sequence Recognition*. PhD thesis, McGill University, (Computer Science), Montreal, Qc., Canada.
- Bengio, Y. (2009). *Learning Deep Architectures for AI*. *Foundations and Trends in Machine Learning*, V2(1). Now Publishers.
- Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(8):1798–1828.
- Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H. (2007). Greedy layer-wise training of deep networks. In Cowan, J. D., Tesauro, G., and Alspector, J., editors, *Advances in Neural Information Processing Systems 19 (NIPS)*, pages 153–160. MIT Press.
- Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166.

Transparência 13: Deep Neural Networks em Sistemas Híbridos (1/4)

Seguindo a tradição de outros paradigmas de Aprendizado de Máquinas (ML), *deep learning* também adota **abordagens não convencionais** para melhorar sua performance. Por exemplo, as abordagens de *CNNs* e *DBNs* podem ser combinadas para formar as *Convolutional Deep Belief Networks (CDBNs)*, que alcançaram o estado da arte. Outras abordagens como *deep learning* utilizando *Linear Support Vector Machine*, *Deep Q Networks*, *Deep Kernel Machines*, *Deep Stacking Networks*, e etc. provaram suas eficiências melhores que as abordagens Tradicionais.

Combinar dois ou mais paradigmas em Aprendizado de Máquinas para alcançar melhora na performance dos algoritmos considerados não é um conceito novo. As **abordagens não convencionais** envolvendo *Deep Learning* podem então serem classificadas em três categorias principais:

- 1) *Deep learning* utilizando arquiteturas convencionas, tais como, *CNNs*, *DBNs* e etc. para formar, por exemplo, as *CDBNs* (**Deep Learning do Tipo 1**);
- 2) *Deep Learning* utilizando outros paradigmas de aprendizagem de Máquinas, tais como, Aprendizado por Reforço (*Reinforcement Learning, RL*), *SVMs* e etc (**Deep Learning do Tipo 2**).
- 3) *Deep Learning* utilizando princípios de Computação Evolutiva (*Evolutionary Computation, EC*) (**Deep Learning do Tipo 3**).

Transparência 14: Deep Neural Networks em Sistemas Híbridos (2/4)

Exemplos de *Deep Learning* do Tipo 1:

- Combinação de *CNNs* com *DBNs* que é chamada de *Convolutional Deep Belief Networks (CDBN)* (Lee et al, 2009; Krizhevsky, 2012).
- *Deep Kernel Machine* (Cho e Saul, 2009).
- *Deep Coding Network* (Lin et al., 2010).
- *Spike-and-Slab RBMs (RBMs:Restricted Boltzmann Machines)* (Luo et al., 2012).
- *Deep Stacking Networks* (Hutchinson et al., 2013).
- *Deep Recurrent Neural Networks* (Pascanu, 2013).

Transparência 15: Deep Neural Networks em Sistemas Híbridos (3/4)

Exemplos de *Deep Learning* do Tipo 2:

- Abordagens de Aprendizado de Máquinas com Aprendizado por Reforço (*Deep Q Networks*) (Mnih et al., 2015).
- *Deep Support Vector Machines* (Tang, 2013).

Transparência 16: Deep Neural Networks em Sistemas Híbridos (4/4)

Exemplos de *Deep Learning* do Tipo 3:

Por outro lado, técnicas de Computação Evolutiva são completamente adaptativas para muitas abordagens de aprendizagem de máquinas. Obviamente, isto não seria diferente com *Deep Learning*. Algoritmos Evolutivos são utilizados para otimizar *Deep Neural Networks (DNNs)* e de maneira convencional como nas redes neurais usuais. O sucesso de implementação de algoritmos evolucionários para *DNNs*, para melhora na performance de aprendizagem, é largamente conhecido (Tirumala, 2014).

A utilização de algoritmos genéticos para aprendizado com *Deep Learning* foi proposto pela primeira vez por (Lamos-Sweeney, 2012).

Atualmente a aplicação de *Deep Learning* está limitada ao processamento de linguagem natural, reconhecimento de imagem e outros aspectos relacionados. No entanto, as tendências de pesquisa indicam que, o poder das *DNNs* não está completamente realizado ainda.

Atualmente, são fontes de pesquisa em *DNNs*: mineração de opiniões (Irsoy e Cardie, 2014), classificação de dados taxonômicos (Tirumala e Narayanan, 2015) e análise sentimental de mensagens curtas online como no twitter e facebook (dos Santos e Gatti, 2014).