Other readings

Monday, April 10, 2023 10:21 PM

Deep Learning in a Nutshell: Core Concepts (optional)

This self-contained tutorial by Tim Dettmers covers the key high-level concepts of deep learning and reinforces the basic concepts we covered in the Neural Networks and Deep Learning lectures. There are multiple parts - Part 1 is less technical while Parts 2-4 go into more detail on algorithms.

The link to access Part 1 is here:

https://devblogs.nvidia.com/parallelforall/deep-learning-nutshell-core-concepts/

Deep Learning in a Nutshell: Core Concepts. (2016, September 08). Retrieved May 10, 2017.

From <https://www.coursera.org/learn/python-machine-learning/supplement/afAsX/deep-learning-in-a-nutshell-core-concepts-optional>

Assisting Pathologists in Detecting Cancer with Deep Learning (optional)

This short article is an example of how deep learning is being used in healthcare.

Assisting Pathologists in Detecting Cancer with Deep Learning

Posted by Martin Stumpe (Technical Lead) and Lily Peng (Product Manager), Google Research Blog

From <https://www.coursera.org/learn/python-machine-learning/supplement/9WBdx/assisting-pathologists-in-detecting-cancer-with-deep-learning-optional>

The Treachery of Leakage (optional)

This fun, less-technical read from Colin Fraser reinforces the material in the Data Leakage lecture to provide further explanation and examples on detecting and avoiding data leakage in your machine learning applications.

Here's the link to the article:

https://medium.com/@colin.fraser/the-treachery-of-leakage-56a2d7c4e931

From <https://www.coursera.org/learn/python-machine-learning/supplement/d1qsQ/the-treachery-of-leakage-optional>

Leakage in Data Mining: Formulation, Detection, and Avoidance (optional)

If you want an example in more depth of how data scientists are exploring ways to detect and avoid data

leakage, this technical article proposes one approach: a two-stage process based on "legitimacy tags".

If you're just interested in getting a little more background on the problem along with interesting examples, Sections 1 and 2 (Introduction and Related Work) are also useful to read on their own.

Kaufman, S., Rosset, S., & Perlich, C. (2011). Leakage in data mining. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11*. doi:10.1145/2020408.2020496

From <https://www.coursera.org/learn/python-machine-learning/supplement/P0buA/leakage-in-data-mining-formulation-detection-and-avoidance-optional>

Data Leakage Example: The ICML 2013 Whale Challenge (optional)

In 2013 a machine learning competition offered a prize for the most accurate detection of right whale calls based on audio data. The organizers soon discovered data leakage problems in the first release of the dataset, and this article explains what happened. It's a short but interesting article that serves as an excellent example of how subtle or not-so-subtle leakage can occur in specific features.

https://www.kaggle.com/c/the-icml-2013-whale-challenge-right-whale-redux/discussion/4865#25839#post25839

From <https://www.coursera.org/learn/python-machine-learning/supplement/BFAwf/data-leakage-example-the-icml-2013-whale-challenge-optional>

Rules of Machine Learning: Best Practices for ML Engineering (optional)

This optional reading is intended mainly for software engineers who want to build and deploy machine learning applications in production - especially at scale. The only background knowledge required are the basic machine learning concepts we've covered so far in this course. Written by Google's Dr. Martin Zinkevich, it walks through a set of software engineering best practices for designing and deploying machine learning in software systems - based on years of practical experience at Google.

http://martin.zinkevich.org/rules_of_ml/rules_of_ml.pdf

From <https://www.coursera.org/learn/python-machine-learning/supplement/EoTru/rules-of-machine-learning-best-practices-for-ml-engineering-optional>