

## 0.1 Theano

Theano is a numerical computation library for Python. In Theano, computations are expressed using a NumPy syntax and compiled to run efficiently on either CPU or GPU architectures. Theano is an open source project, primarily developed by a machine learning group at the Université de Montréal [10]. Theano is a Python library that allows developers to define, optimize, and evaluate mathematical expressions involving multi-dimensional arrays efficiently. Some of the features of Theano include tight integration with NumPy, transparent use of a GPU to perform data intensive computations, dynamic C code generation for evaluation expressions faster and support for extensive unit testing and self-verification [7].

## 0.2 Apache Mahout

Apache Mahout, an Apache Software Foundation project, is a distributed linear algebra framework and mathematically expressive Scala DSL designed to aid mathematicians, statisticians and data scientists in implementing their own algorithms [2]. The project was initially driven by Ng et al.'s paper "Map-Reduce for Machine Learning on Multicore" and has since evolved to cover much broader machine-learning approaches [3]. While Apache Spark is recommended back end and core algorithms are implemented on top of Apache Hadoop, it can be extended to other back ends and standalone implementations as well. Apache mahout's core algorithms primarily contain implementations for clustering, classification, and Collaborative filtering [8].

## 0.3 Apache PredictionIO

Apache PredictionIO is an open source machine learning stack for building, evaluating and deploying engines with machine learning algorithms. An open source Machine Learning Server built on top of an open source stack allows developers and data scientists to create predictive engines for any machine learning task. It allows developers to quickly build and deploy an engine as a web service and unify data from multiple platforms in batch or in real-time for comprehensive predictive analytics. It supports machine learning and data processing libraries such as Spark MLlib and OpenNLP [1].

## 0.4 Kubernetes

Kubernetes is an open-source platform designed to automate deploying, scaling, and operating application containers. The name Kubernetes originates from Greek, meaning helmsman or pilot, and is the root of governor and cybernetic. Kubernetes can schedule and run application containers on clusters of physical or virtual machines. However, Kubernetes also allows developers to 'cut the cord' to physical and virtual machines, moving from a host-centric infrastructure to a container-centric infrastructure, which provides the full advantages and benefits inherent to containers [4]. Kubernetes defines a set of building blocks ("primitives") which collectively provide mechanisms for deploying, maintaining, and scaling applications. The components which make up Kubernetes are designed to be loosely coupled and extensible so that it can meet a wide variety of different workloads [9].

## 0.5 Microsoft Cognitive Toolkit

The Microsoft Cognitive Toolkit (CNTK) is an open-source toolkit for commercial-grade distributed deep learning. CNTK allows the user to easily realize and combine popular model types such as feed-forward DNNs, convolutional neural networks (CNNs) and recurrent neural networks (RNNs/LSTMs). CNTK implements stochastic gradient descent learning with automatic differentiation and parallelization across multiple GPUs and servers [6]. CNTK currently supports Windows and Linux platforms. CNTK also supports the description of neural networks via C++, Network Definition Language (NDL) and other descriptive languages such as Python and C# [5].

### References

- [1] Apache Software Foundation. *Apache PredictionIO*. Web. URL: <https://predictionio.apache.org/#what-is-apache-predictionio> (cited on page 1).
- [2] Apache Software Foundation. *Mahout*. Web. Dec. 2017. URL: <https://mahout.apache.org> (cited on page 1).
- [3] Grant Ingersoll. *Introducing Apache Mahout*. Web. Sept. 2009. URL: <https://www.ibm.com/developerworks/java/library/j-mahout/> (cited on page 1).
- [4] Kubernetes. *Kubernetes*. Web. URL: <https://kubernetes.io/> (cited on page 1).
- [5] Matthew Mayo. *Microsoft Deep Learning Brings Innovative Features – CNTK Shows Promise*. Web. Feb. 2016. URL: <https://www.kdnuggets.com/2016/02/microsoft-deep-learning-brings-innovative-features.html> (cited on page 2).
- [6] Microsoft. *The Microsoft Cognitive Toolkit*. Web. URL: <https://www.microsoft.com/en-us/cognitive-toolkit/> (cited on page 2).
- [7] Theano. *theano*. Web. URL: <http://deeplearning.net/software/theano/> (cited on page 1).
- [8] Wikipedia. *Apache\_Mahout*. Web. Sept. 2017. URL: [https://en.wikipedia.org/wiki/Apache\\_Mahout](https://en.wikipedia.org/wiki/Apache_Mahout) (cited on page 1).
- [9] Wikipedia. *Kubernetes*. Web. Feb. 2018. URL: <https://en.wikipedia.org/wiki/Kubernetes> (cited on page 1).
- [10] Wikipedia. *Theano*. Web. Feb. 2018. URL: [https://en.wikipedia.org/wiki/Theano\\_\(software\)](https://en.wikipedia.org/wiki/Theano_(software)) (cited on page 1).