Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 119 (2015) 479 – 485
Procedia Engineering
www.elsevier.com/locate/procedia

13th Computer Control for Water Industry Conference, CCWI 2015

Applications of deep learning for smart water networks

Zheng Yi Wu[a]* , Mahmoud El-Maghraby[b] and Sudipta Pathak[b]

[a]*Applied Research, Bentley Systems, Incorporated, Watertown CT, USA*
[b]*Department of Computer Science and Engineering, University of Conneticut, Storrs, CT, USA*

Abstract

Deep Learning (DL) is the state-of-art paradigm of Artificial Neural Network (ANN) computing. It is a new breakthrough in machine learning, and differentiates from the conventional or shallow learning algorithms by emulating the six-layer human neocortex, which is unique for human brain containing billions of interconnected neurons. Unlike canonical ANN, DL is capable of self-learning data features by mimicking the self-learning functions layer by layer in human cortex and creating a data-driven model with the given dataset. This paper reports the initial applications of deep learning for simulation, optimization and operation control of water distribution systems. It elaborates the development of efficient deep learning framework with potential applications of facilitating the data fusion, system simulation and predictive analysis, detection of abnormal events from the recorded time series data (pressures, flows and consumptions etc.), water usage prediction, construction of a meta-model as a surrogate to the physics-based models (hydraulic and water quality), and acceleration of the solution search for smart water distribution management, which aims at improving operation efficiency, reducing carbon footprint, and exceling customers' expectation.

Peer-review under responsibility of the Scientific Committee of CCWI 2015

Keywords: smart water netwroks, modeling, big data, data analytics, data-driven model, deep learning

1. Introduction

Provision of sufficient water quantity and water quality is one of the key challenges, among others, which the cities around the world are facing in sustaining economic growth with good/health living conditions. Public water

* Corresponding author. Tel.: +1-203-8050562; fax: +0-000-000-0000 .
E-mail address: zheng.wu@bentley.com

1877-7058
Peer-review under responsibility of the Scientific Committee of CCWI 2015
doi:10.1016/j.proeng.2015.08.870

utilities and private water companies are under increasing pressure in addressing the challenge. They are expected to look for and adopt the new technologies and solutions so that the efficiency of the total expenditure (Totex) can be maximized with the best outcome of the investment in water systems. Smart water network is emerging as an integrated approach to enable a city to improve the Totex efficiency for its water distribution system.

Making a water network truly smart is not only a daunting task, but also represents great opportunities for innovations and new business of smart metering, monitoring, data acquisition, data management, and advanced analysis, communication and automation control. There are different interpretations or definitions of smart water networks. Hardware vendors, such as meter and sensor companies, often define smart water networks as a fully integrated products with emphasis on hardware-related or based solutions [1], which (1) enable remote and continuous monitoring, diagnosing potential problems; (2) comply regulatory requirements for water quality and conservation; and (3) provide customers with consumption information about their behaviours on water usage. Software vendors [2] [3] tend to emphasize on how software solutions can help to make a water network smart or intelligent by operating the system using real-time or near real-time information. The features of a smart water network include monitoring/sensing with instrumentation, data management, data analytics for useful/actionable information retrieve or extraction, systematic analytics including simulation and optimization modelling for decision making, and finally the automation control for triggering/communicating the instruments in the field. A truly smart water network needs to be ‘smart’ at each of the steps to achieve the best outcomes of water network management and operation.

In this paper, we report the preliminary investigation of deep learning, one of the breakthrough in artificial intelligence, for efficiently and effectively extracting the intelligence from data, with potential of wide applications in simulation and prediction analysis that serves as the data-driven engine of smart water network management.

2. Deep learning

Deep Learning (DL) has been recognized as one of ten breakthrough technologies in 2013 according to MIT Technology Review [4]. Deep learning was described as the technology that allows AI to finally getting smart. Deep learning is a set of algorithms in machine learning that attempts to model high-level abstractions in data with multiple non-linear transformations. It is to mimic the activity in many layers of neurons in the brain's cortex. For instance, recognition pathway in the visual cortex has multiple stages, with lots of intermediate representations. As illustrated in Fig. 1, after the scene been intercepted in the retina, it paths multiple stages, for example, after V1 stage, the scene is interpreted in simple visual forms like edges and corners. Stage V4 groups such representations into groups of features. Finally, stage AIT produces high-level descriptions as faces and objects. So, each stage gives a higher abstraction of its inputs.

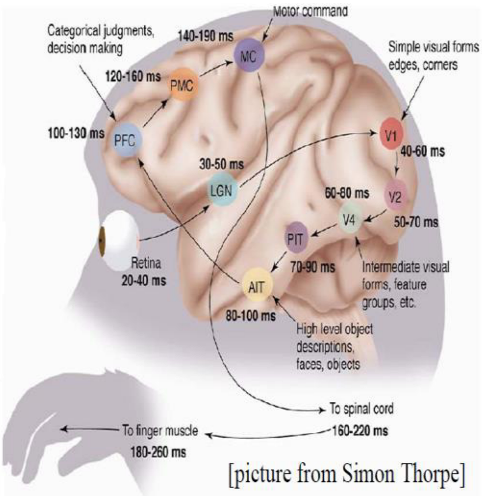

Figure 1. Information path for human visual recognition

DL has been successfully applied to many commercial products [5]. Some of those applications are object recognition, speech recognition, language translation, and self-driving cars. In 2011, Google started the Google Brain project as one of the Google X projects. Google built the software that uses the deep learning technique to analyse 10 million photos taken from YouTube videos and learned to recognize thousands of objects, including human and cat faces, without human guidance. The project's technology is currently used in the Android Operating System's speech recognition system and photo-search for Google+. It enables high-quality speech recognition, practical computer vision, email spam blocking. Later, Microsoft challenges Google's Artificial Brain with a project named 'Project Adam'. With project Adam, Microsoft created one of the best photograph classifier. They used the largest data set they can find, train massive neural network with over 2 billion connections, with 30 times fewer machines and develop image classifier that produces double the accuracy of other systems in industry at that time. Facebook was not an exception. Facebook used the Deep Learning technique to help understanding their users and there data better. Netflix currently is aiming to build an online recommendation engine that outperforms even your closest friends, using the Deep Learning technique. Finally, Chinese internet search company Baidu has also adopted DL technology for improving their existing products and developing new ones especially for visual search.

The most advantage feature of the deep learning is automatic feature extraction. The traditional model for pattern recognition is to apply hand-crafted feature extractor to the images before applying the trainable classifier. In contrast, the deep learning systems are able to automatically extract the features by training the system with unlabelled data, and then use those automatically extracted features to classify the image using a trainable classifier.

The deep learning architecture is consisting of multiple layers. After applying the input to the system, each layer can be individually trained to extract a higher level of abstractions from its inputs. For example, if the input in a large set of unlabelled images, the output of the first layer (after training) are the features of edges and corners. This output is used to train the next layer to be able to produce group of features, such as body parts, noses, eyes, and so on. The third layer (after being trained) can extract a higher level of abstractions, and recognize the desired image, e.g. a human face. This process shows the ability of the Deep Learning system to automatically extract features from the inputs. The advantages of Deep Learning are summarized below:

- It is representationally efficient than the shallow artificial neural networks. It requires less computational units than the shallow artificial neural networks to perform the same function.
- It automatically learn (or extract) features without user's interaction or guidance.
- In many cases as published, deep learning gives better accuracy than its counterpart conventional artificial neural networks.
- Deep learning is able to use the available large amount of unlabeled data.

Some of the commonly employed deep Learning architectures are Deep Belief Networks (DBN), Stacked Auto-encoders (SAE), and Convolutional Networks (CN). DBN is chosen for our applications, as in many publications it outperforms the SAE in terms of accuracy. Also DBN is more suitable for non-computer vision applications than convolutional neural networks that is particularly developed for computer vision applications.

3. DBN framework development

Hinton, Osindero and Teh [6] have proposed Deep Belief Network (DBN), and showed that it is possible to stack a number of Restricted Boltzmann Machines (RBMs) and train them in a greedy layer wise manner to learn complex models much better than conventional ANN architectures. A DBN architecture can be thought of a single layer RBM. An RBM is a fully connected bipartite graph with two layers namely visible and hidden layers [7]. The visible and hidden layers are comprised of neurons capable of working with binary or Gaussian values. Based on that the architecture is named as binary-binary, Gaussian-binary or Gaussian-Gaussian RBM. Fig.1 shows an RBM with four visible neurons and three hidden neurons.

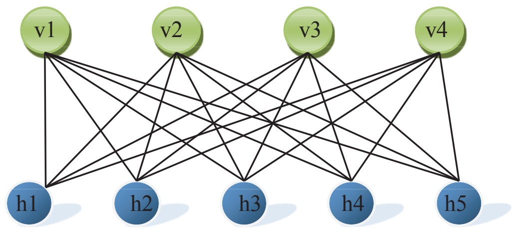

Fig. 2. Restricted Boltzmann machine

DBN models conditional probability distribution between observed vector and hidden layers as follows:

$$P(x, h^1, h^2, h^3, \dots \dots \dots \dots, h^l) = \left(\prod_{k=0}^{l-2} P(h^k|h^{k+1})\right) P(h^{l-1}, h^l) \quad (1)$$

where x corresponds to the input layer and is equivalent to h^0. $P(h^{k-1}|h^k)$ is the joint probability distribution of visible units of a RBM at layer k conditioned on the hidden units of the RBM at layer k. Fig. 2 illustrates the DBN architecture.

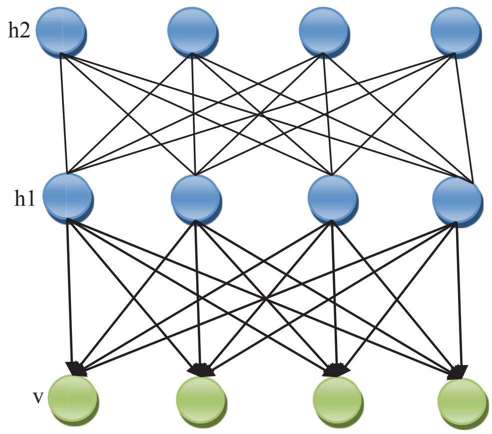

Fig. 3. Deep Belief Network

For solving a classification problem, one can stack layers of binary-binary RBMs for unsupervised pre-training. Using a stack of binary-binary RBMs for predictive analysis is no effective. A Gaussian-binary RBM must be implemented with the following energy function [8], given as:

$$E(v, h) = -\sum_{i \epsilon visible} \frac{(v_i - a_i)^2}{2{\sigma_i}^2} - \sum_{j \in hidden} b_j h_j - \sum_{i,j} \frac{v_i}{\sigma_i} h_j w_{ij} \quad (2)$$

Where v_i and $h_j$are the states of the visible and hidden units respectively. w_{ij} is the weights from neuron i to neuron j. σ_i is the standard deviation of the Gaussian noise for visible unit i. Learning the noise associated with each visible unit poses a challenge, but it can be overcome by normalizing each component of the data by subtracting the mean and divided by the variance. The learning rate is adjusted by reducing it by one or two magnitude from what is used for binary-binary RBM training. Reduction in learning rate is required to keep the weights emanating from certain components becoming very large. We employed Gaussian visible units with rectified linear hidden units for implementing Gaussian-binary RBM [8]. In this model the hidden units are approximated by $\max(0 + N(0,1))$ where $N(0,1)$ is a gaussian noise with zero mean and unit variance. Training DBN comprises of an unsupervised pre-training followed by a fine tuning [9]. The training process is outlined as follows.

- Firstly, train RBM layer using the original/scaled data as input to its visible layer. The first layer of RBM is executed for a sufficiently large number of epochs to obtain a representation of the input that is used by the second RBM layer's visible units.
- Secondly, train the second layer of RBM using the output of the first RBM layer as input.
- Thirdly, each subsequent layer is trained in the same manner to find good initialization parameters for the network.
- Finally, fine tune the parameters of the network using a supervised training algorithm.

4. Applications

In the previously published works [10-12], a GPU-based ANN modelling tool was developed to train a surrogate model for the extended period simulation (EPS) of a water distribution system. The training was done on GPU to reduce the training time from more than 10 hours using CPU alone to a few minutes by using GUP computing. A large water system, containing a dozen tanks and several pump stations, was employed as case study. A total of 13 ANNs have been trained, with each producing one output, for representing the whole system operation. In addition, before training the ANNs, sensitivity analysis, much like manual feature extraction for applying ANN for image recognition, must be performed to figure out what inputs are sensitive to the desired output. This is tedious and time consuming. With deep learning approach, e.g. a DBN, it is expected that we can eliminate the requirement for sensitivity analysis and also avoid using multiple ANNs to represent the whole system as shown in Fig. 3.

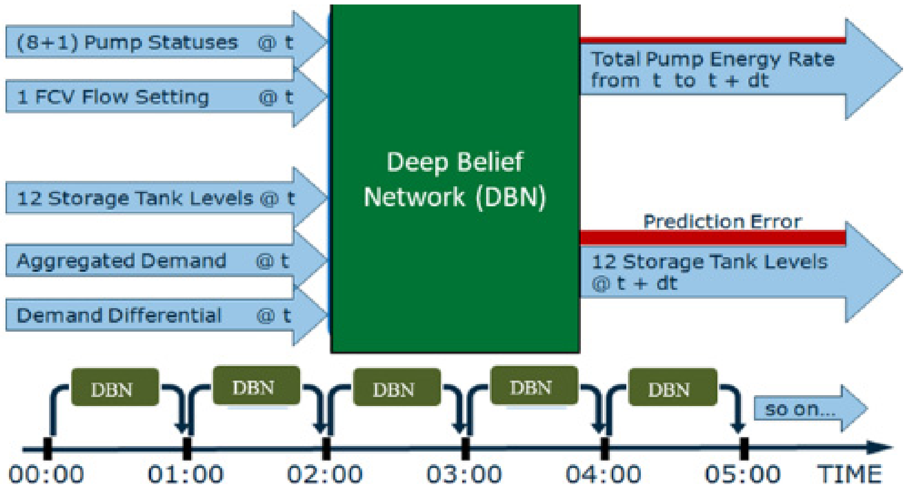

Fig. 4. Configuration of deep learning model structure for EPS simulation

With DBN method, a user starts by preparing all the necessary input files, and the configuration file, which contains all the information, such as the number of hidden layers, the number of neurons in the input layer, the number of neurons in the output layer, the number of neurons in each hidden layer, and the number of epochs for per-layer training, needed to build the DBN. The DBN tool can run in three modes including train, test, or EPS simulation. After training is completed, the weights can be automatically fed into DBN for testing and/or simulation as desired.

Fig. 4 illustrates an example results after applying the DBN as a data-drive model on the water distribution network. It compares the tank levels simulated by using the hydraulic model, the conventional ANN, and the DBN. As shown in the figure, very similar accuracy has been achieved by both ANN and DBN, however, just one DBN mode is trained for the system without sensitivity analysis required for constructing the deep neural network.

Finally, the average error for all tanks after applying 13 ANN models with sensitivity analysis is 0.05156m, while the average error for all tanks using DBN model is about 0.05045m. Deep learning approach has an advantage of eliminating the need for sensitivity analysis without sacrificing the accuracy of the results. A slightly better accuracy is achieved using the deep Learning approach.

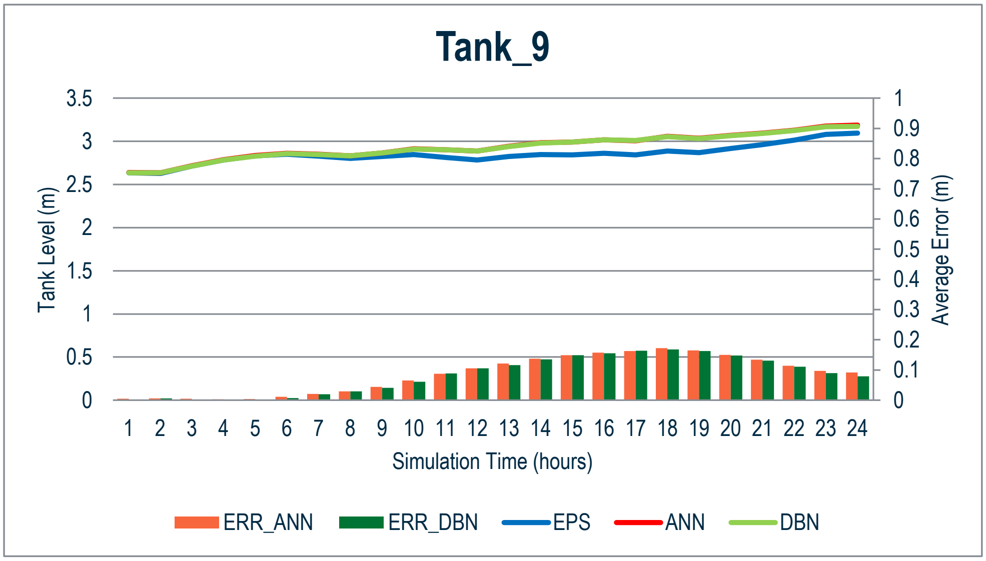

Fig. 5. Example simulation results by trained DBN model

5. Conclusions

A deep learning algorithm DBN has been developed and applied to WDS modelling. Preliminary results show that it outperforms the conventional ANN, enables automatic feature learning with slightly better accuracy. This on-going research project is aiming at developing a generic, robust, effective and efficient data-driven analysis tool by using heterogeneous computing paradigm, e.g. many-core architecture with general purpose GPU, and also extending its application to a wide range of data-driven analysis for the performance modelling of various infrastructure systems.

Reference

[1] Sensus, Water 20/20 Bring Smart Water Networks into Focus. http://sensus.com/smartwaternetworks, (accessed May 25 2015).
[2] SWAN What are Smart Water Networks? http://www.swan-forum.com/swan-documents.html, (accessed May 25 2015).
[3] IBM, Water Efficiency Management. http://www-03.ibm.com/software/products/en/intelligentwater, (accessed May 25 2015).
[4] MIT Technology Review, Ten Breakthrough technologies 2013. http://www.technologyreview.com/lists/breakthrough-technologies/2013/, (accessed May 20 2014).
[5] Y. LeCun, Y. Bengio and G. E. Hinton, Deep Learning, Nature, 28 May 2015, Vol 521, pp436-444.

[6] G. E. Hinton, S. Osindero & Y.-W. Teh, A fast learning algorithm for deep belief nets. *Neural Comp.* **18,** 1527–1554, 2006.
[7] G.E. Hinton and R.R. Salakhutdinov, Reducing the Dimensionality of Data with Neural Networks. Science, 28 July 2006, Vol. 313. no. 5786, pp. 504 – 507.
[8] V. Nair and G. E. Hinton, Rectified linear units improve restricted Boltzmann machines. In Proc. 27th International Conference on Machine Learning, Haifa, Israel, June 21-24, 2010, http://www.icml2010.org/, (accessed at May 15, 2015).
[9] A Practical Guide to Training Restricted Boltzmann Machines, accessed at https://www.cs.toronto.edu/~hinton/absps/guideTR.pdf, May 20 2015.
[10] Z. Y. Wu & A. A. Eftekharian, Parallel artificial neural network using CUDA-enabled GPU for extracting hydraulic domain knowledge of large water distribution systems. *ASCE Annual World Environmental and Water Resources Congress*. Palm Springs, CA, USA, 2011.
[11] M. Behandish & Z. Y. Wu, GPU-based ANN configuration and training for water distribution system analysis. *ASCE Annual World Environmental and Water Resources Congress*. Albuquerque, NM, USA, 2012.
[12] Z. Y. Wu and M. Elmagraby, Portable GPU-based Artificial Neural Networks for Accelerated data-driven Modelling. 11th International Hydroinformatics, New York City, NY, USA, 2014.