A
PROJECT REPORT
ON

Prism: Graphics Accelerated Data Analysis using Deep Neural Networks

by

Parth Sane	113A1079
Santosh Shettar	113A1083
Amish Thakkar	113A1091

Under the guidance of
Prof. Ravindra Agarwal

Submitted in partial fulfillment for the degree of

BACHELOR OF ENGINEERING
in
Computer Engineering

DEPARTMENT OF COMPUTER ENGINEERING
SIES GRADUATE SCHOOL OF TECHNOLOGY

NERUL, NAVI MUMBAI-400 012
UNIVERSITY OF MUMBAI

ACADEMIC YEAR 2016-17

CERTIFICATE

This is to certify that project titled "Prism: Graphics Accelerated Data Analysis using Deep Neural Networks" is a bona-fide record of Project-B carried out by the following students of final year in Computer Engineering.

Sr. No.	Name here	Roll no here	Exam No.
1	Parth Sane	113A1079	
2	Santosh Shettar	113A1083	
3	Amish Thakkar	113A1091	

This project is carried out for the partial fulfillment of requirements for the degree of Bachelor of Engineering (B.E.) in Computer from University of Mumbai, Mumbai for academic year 2016-2017.

Prof. Ravindra Agarwal	Dr. Rizwana Shaikh	Dr. Vikram Patil
Guide (Internal)	Head of Department (Electronics and Telecommunication)	Principal (S.I.E.S. GST)

We have examined this report as per university requirements at SIES Graduate School of Technology, Nerul, Navi Mumbai on ____________________

Signature: Signature:

Name: Name:

(Examiner-I) (Examiner-II)

Acknowledgments

We would thank Prof. Agrawal for his guidance in developing the correct thought process as our project advisor. Special thanks and regards to our esteemed principal and HOD for supporting us through the project. Finally, last but the most, Parth's parents for sponsoring his PC and enabling us to complete this experiment and for their unwavering support.

Abstract

When a customer of say a corporate telecom major like Airtel decreases spending on his or her number, they tend to provide more offers to the customer. Now this logic may sound straightway outrageous but this is true based on churn models. Churning refers to customers leaving a service and switching to another provider. The standard industry practice is to generate churn models for every consumer, so that they can provide business intelligence to increase customer retention and spending. Thus this translates to greater profit for a company. Although such processes can be carried out with a GPU, the software that exists like SAP handles everything on the CPU. This leads to longer delivery period for business intelligence and analytics which in turn means less time spent doing business and more time spent waiting. GPUs can generate huge number of threads and thus introduce parallelism in this area. In addition to the above advantage the churn models have assisted learning and cannot learn on its own unlike Deep Neural Networks which apply the concept of machine learning. This project aims to create a Windows based general purpose application to try predicting churn with better analytics. The scope of the project is limited to test data sets freely available on the internet as real data can only be obtained from corporates who are rather unlikely to share sensitive data with non-employees. Thus we aim to create a churn model that uses GPU based DNNs which not only will improve execution time for large data sets, it will also prove to be a huge plus point for corporate customers who want to use this software.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

The concept of our project is to build a churn model which helps in predicting whether a user may churn or not i.e the user changes or leaves the existing service provider. Project is implemented in Matlab as it allows for rapid prototyping. Churn modelling requires computing of large datasets. Our project mainly focuses on computing these large data sets with the help of GPU acceleration.

Our technique lies in the paper presented by Sane and Agrawal [8][8]. We convert the textual data into an image using Min-Max normalisation as described in the paper cited above. This conversion is an intermediate step which helps us feed data into the neural network. As mentioned the main highlight exists in preprocessing the data for predictive analysis [1].

1.1 Background

As we can see in [1] churn ananlysis and prediction is a big part of managing customer relationships. The idea behind this project was to fundamentally increase the efficiency of this analysis process by introducing the concept of Prism. The concept of Prism is to split the data and perform analysis on it parallely on a Graphics Processing Unit (GPU). This decreases the time spent on analysing and executing neural networks by putting them on a GPU.

1.2 Objective

Objective of the project is to demonstrate a significant speedup in execution time of analysis of data sets on Neural Networks.

1.3 Purpose, Scope and applicability

Purpose, Scope and Applicability are important parameters to be discussed to give a definite direction to the project. They are described briefly as follows:

1.3.1 Purpose

The project aims to provide a way to analyse your dataset faster than ever before by cutting down time required by graphics accerelation via GPU compute technology.

1.3.2 Scope

The scope of the project is to clean,extract and transform datasets to implement on a Neural Network.

1.3.3 Applicability

Applicabilty of done work is universal in the sense that data is independent of language. As long as the dataset under consideration is numeric, the technique is applicable.

Chapter 2

LITERATURE SURVEY

2.1 Literature Survey

The highlight of the paper is the idea of using the normalized data as an image and using the generated image as input to the neural network. This concept was inspired by the paper on churn data model viewed as periods of time [2].

2.1.1 Dataset Used for the Experiment

Our experiment required data to be processed and used as an input. This was obtained online from bigml [5]. We, thus, obtained a labelled dataset, which enables us to use supervised learning approach for our neural network. Here, our sample dataset appears to be monthly rather than weekly as compared to the dataset used by Wangperpong et al. [4]. Our sample dataset consists of rows - customers and columns as attributes associated with customers. The attributes, which are significant, are not identified in our dataset, but we only consider numeric values while trying to predict the churn rate.

2.1.2 Pattern Matching and Clustering Neural Networks Using Supervised Learning

Furthermore, there has been a lot of work done in pattern matching via supervised learning. Schwenker et al. [3] describe this concept very clearly in their detailed analysis of the topic. This approach to supervised learning is heavily applied to neural networks [7], [8]. Neural networks thus find far and wide reaching applications that can have major impacts on society like the traffic camera system in the reference [9].

2.2 Problem Definition

Customer retention is a huge requirement for any service based business. This is especially more emphasised in the Telecommunications industry where cellular carriers must retain customers to remain profitable or risk losing money. Churn analysis aims to predict the churn which is the act of a consumer leaving or switching service. Thus these predictions offer a chance to act on potential loss of subscribers. But this deep analysis of data using Neural Networks requires immense computing power and time both of which are critical in a subscription model of business. The problem is to create a technique to demonstrate speedup in the analysis process of churn data and then provide predictions with accuracy over 90% reliably.

Chapter 3

IMPLEMENTATION

3.1 Hardware

The main system has i7, 6700k processor with 16gb DDR4 RAM clocked at 2133MHz. Graphics Processing Unit (GPU) which is Nvidia Titan X (Maxwell Gen.) 12gb GDDR5 VRAM. This particular system was running Windows 10 Pro x64 build 1607. For the sake of comparison we implemented the project on MacBook Air 13 inch early 2015 edition with specs of intel core i5 at 1.6kHz and 4gb DDR 3 1600MHz. This particular system did not possess a discrete GPU. Therefore it provided a good benchmark analysis of a graphics enabled version vs CPU only version of the project. MacBook Air was running MacOS sierra as the operating system.

3.2 Software

Matlab was chosen as the software to build the project. Dealing with large amounts of data meant that we needed a quick way to evaluate results and import and export data conveniently. All these features were readily made available by Matlab and hence it was the ideal software for completing the project.The GPU acceleration provided by Matlab is through the Cuda runtime which is transparently available to the programs with only slight modifications required for GPU based code.
The following steps are taken in software to execute the project:

1. Import data from the data source.

2. Run transformation of the dataset to an image.

Table 3.1: SYSTEM CONFIGURATION FOR EXPERMIENTAL SYSTEMS

Processor	Physical RAM	CUDA GPU name	MATLAB GPU support	Operating system and build
Intel Core i7 6700k @ 4 GHz	2 8 GB Kingston DDR4 Non ECC @ 2133 MHz	NVIDIA Titan X Maxwell Architecture 12 GB GDDR5	Yes	Windows 10 64 bit Build 1607, Custom PC Build
Intel Core i5 @ 1.6 GHz	4 GB DDR3 Non ECC @ 1600 MHz	N/A	No	MacOS Sierra 10.12.2 Build 16C67 on MacBook Air Early 2015, 13 inch

3. Feed generated image into the neural network generated by Matlab.

The results of the above described process is given in the next section.

Chapter 4

RESULT AND DISCUSSION

The implementation described in the previous section leads to a lot of results and requires explanation.

- The project is mainly to improve execution speed of analysis of datasets. It deals with converting the datasets from its numerical form to the 8 bit grayscale colour domain

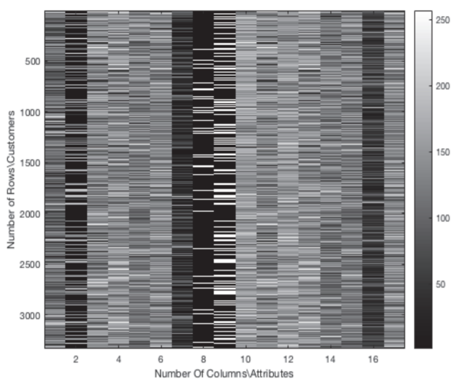

Figure 4.1: Customer records and attributes as an image.

The neural network is implemented with the Neural Network Toolkit (NNT) [11] in MATLAB [10]. The NNT contains predefined types of neural networks for clustering, fitting, pattern recognition, and time series. These types make it possible to instantly deploy the neural networks, which otherwise would take considerable time to set up. The NNT handles all initializing of weights and other trivial processes. We use the pattern recognition pre-set for the

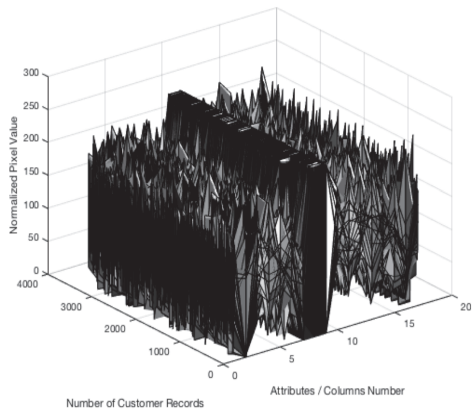

Figure 4.2: 3D surface plot of customer records and attributes.

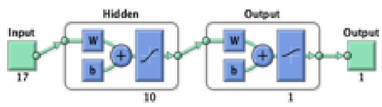

Figure 4.3: Structure of neural network.

neural network so that we can detect trends and patterns via neural networks. As we mentioned earlier, we are using the supervised learning approach since we have a labelled dataset for churn analysis. For creating the neural network, we have to select certain parameters such as < parameter x, parameter y >. We have to select the number of hidden layers. This can have an impact on the accuracy of prediction that is obtained. We use 10 hidden layers, which is the default number of layers. The selected pre-set pattern recognition neural network in the neural network is implemented with the Neural Network Toolkit (NNT) [11] in MATLAB [10]. The NNT contains predefined types of neural networks for clustering, fitting, pattern recognition, and time series. These templates make it possible to instantly deploy neural networks, which otherwise would take considerable time to set up. The [11] handles all initialization of weights and other trivial processes. are handled by the NNT[11]. NNT[11] has to be trained like any other network for creating a model. We use Scaled conjugate gradient back propagation for supervised training [6]. This is the default algorithm for the pattern recognition neural network in the NNT. The neural network as illustrated in Fig. 3 has 17 inputs, since the input customer attributes are 17. It gives a binary output as evident in 4.3. We observed varying accuracies since the weights are all initialized randomly. Anybody wishing to recreate the experiment may not get exact matching results due to this randomness. But, the accuracy of

Table 4.1: EXECUTION TIME ON EXPERIMENTAL SYSTEMS

Computer description	CPU execution time (seconds)	GPU execution time (seconds)
High End Custom PC	0.001553	0.001051
MacBook Air Early 2015	0.004755	N/A

prediction should be nearby in the neighbourhood of 12% or more [8].

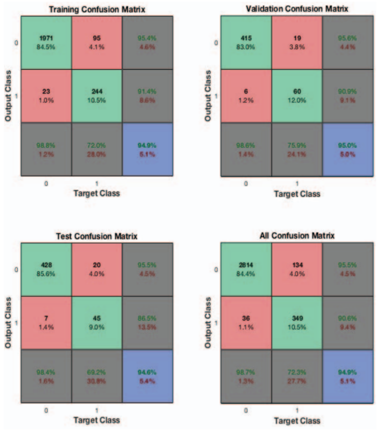

Figure 4.4: Confusion plot for neural network.

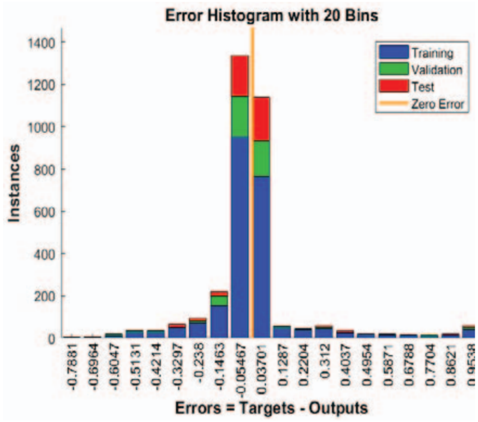

Figure 4.5: Error histogram for the neural network

Hence we obtained 94.9% accuracy which can be seen in the confusion plot in 4.4.

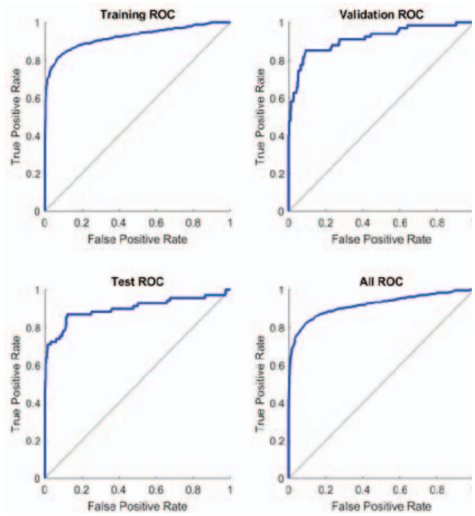

Figure 4.6: Receiver operating characteristic (ROC) curves.

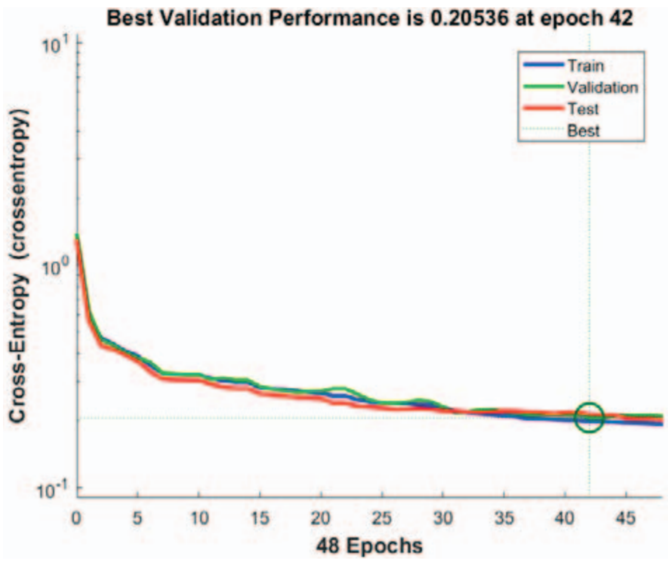

Figure 4.7: Performance plot for the neural network.

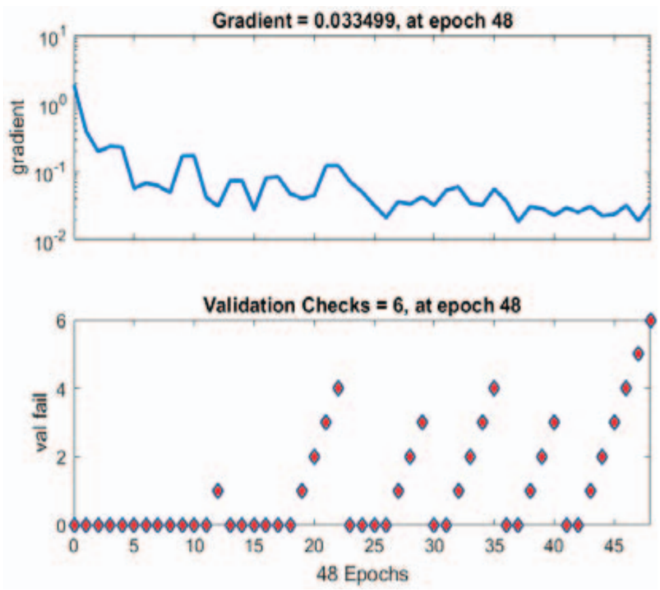

Figure 4.8: Training state for the neural network.

Chapter 5

CONCLUSION AND FUTURE SCOPE

5.1 Conclusion

The project's concept mainly includes the prediction of churn in telecom industry. To acheive the main goal of this project, we had planned to use GPU acceleration on large datasets and convert the textual data into an image format data. The idea of this project is implemented in MATLAB.The converted dataset is then fed into the Neural Network which is built using the feature of MATLAB's Neural Network Toolkit.
We effectively developed a better way to pre-process data so as to improve performance of processing data for the GPU.

5.2 Future Scope

This definitely will give better results as compared to execution on a CPU. Training is many times faster on a GPU as compared to a CPU. This is evidenced in many papers published earlier. Future scope and work can be expanded to actually implementing the Neural Network on the GPU.

References

[1] Li, Hui, Deliang Yang, Lingling Yang, and Xiaola Lin,*Supervised Massive Data Analysis for Telecommunication Customer Churn Prediction*,Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom)(BDCloud-SocialCom-SustainCom),pp. 163-169, IEEE, 2016.

[2] A. Wangperawong, C. Brun, O. Laudy, and R. Pavasuthipaisit, *Churn analysis using deep convolutional neural networks and autoencoders,*,2016,[Online]. Available: https://arxiv.org/abs/1604.05377. Accessed: Jan. 25, 2017.

[3] F. Schwenker and E. Trentin, *Pattern classification and clustering: A review of partially supervised learning approaches*, Pattern Recognition Letters , vol. 37 ,pp. 414, Feb. 2014.

[4] Amezcua, Jonathan, P. Melin, and O. Castillo, *A Neural Network with a Learning Vector Quantization Algorithm for Multiclass Classification Using a Modular Approachs*, Recent Developments and New Direction in Soft-Computing Foundations and Applications, Springer International Publishing, pp.171184, 2016.

[5] B. Carlo Migel and P. S, *Convolutional neural network for vehicle detection in low resolution traffic videos* in Region 10 Symposium (TENSYMP), 2016 IEEE, IEEE, 2016.

[6] M. F. Mller, *A scaled conjugate gradient algorithm for fast supervised learning*, Neural Networks, vol. 6, no. 4, pp. 525533, Jan. 1993.

[7] V. Sowmya, D. Govind and K. Soman, *Significance of incorporating chrominance information for effective color-to-grayscale image conversion*,Signal, Image and Video Processing, vol. 11, no. 1, pp. 129136, 2016.

[8] Parth Sane and Ravindra Agrawal, *Pixel Normalization from Numeric Data as Input to Neural Networks*,IEEE WiSPNET 2017 conference,pp. 2250-2254,2017.

[9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, *ImageNet: A large-scale hierarchical image database,*, 2009 IEEE Conference on Computer Vision and Pattern Recognition ,2009. pdf. Accessed: Jan. 25, 2017.

[10] MATLAB 2017a, The MathWorks, Natick, 2017.

[11] T. MathWorks, Neural network Toolbox - MATLAB, in Mathworks.com, 2016. [Online]. Available: https://www.mathworks.com/products/neural-network.html. Accessed: Jan. 25, 2017.