

Automatic Selection of Open Source Multimedia Softwares Using Error Back-Propagation Neural Network

Deepika¹, Suman², Arshdeep Singh³, Dr. G.N. Verma⁴

¹M.Tech Student Sri Sukhmani Institute of Engineering & Technology, Punjab, India

²Assistant Professor Sri Sukhmani Institute of Engineering & Technology, Punjab, India

³M.E Scholar National Institute of Teachers Training & Research, Chandigarh, India

⁴Coordinator Sri Sukhmani Institute of Engineering & Technology, Punjab, India

Abstract—

Open source opens a new era to provide license of the software for the user at free of cost which is advantage over paid licensed software. In Multimedia applications there are many versions of software are available and there is a problem for the user to select compatible software for their own system. Most of the time while surfing for software a huge list of software opens in response. The selection of particular software which is pretty suitable for the system from a real big list is the biggest challenge that is faced by the users. This work has been done that focuses on the existing open source software that are widely used and to design an automatic system for selection of particular open source software according to the compatibility of users own system. In this work, error back-propagation based neural network is designed in MATLAB for automatic selection of open source software. The system provides the open source software name after taking the information from user. Regression coefficient of 0.93877 is obtained and the results shown are up to the mark and can be utilized for the fast and effective software search.

Keywords—Back-propagation, Open Source, Neural network

I. Introduction

The motivation to build an automatic software selection system comes from the human brain system. Our brain is able to categorize many complex tasks. Human brain has billions of processing units called neurons. Human eyes can sense the external environment which acts as a receptor. Photons on retina receive the information from external world. Neurons in the brain process the information in massively parallel mode. The output of the brain activates the muscles which is called effector action. Effectors interact with the environment. The biological neural system learns from examples and can take decisions based on purely new environment [1]. Human brain has the excellent power of decision making. But this power is very much limited to local problems. Human can take decision well when the number of decisive class is limited. The open source software exists large in number. The configuration of computer system changes the selection of software completely. In these conditions it is very difficult for the humans to decide which software is compatible.

Open source refers to a computer program in which the source code is available to the general public for use and/or modification from its original design. Open-source code is meant to be a collaborative effort, where programmers improve upon the source

code and share the changes within the community In production and development, open source as a development model promotes a universal access via a free license to a product's design or blueprint, and universal redistribution of that design or blueprint, including subsequent improvements to it by anyone.

M. St. Laurent [2] presents Open-source software made available with a license in which the copyright holder provides the rights to study, change, and distribute the software to anyone and for any purpose. Irina Guseva [3] explains the top four reasons (as provided by Open Source Business Conference survey) individuals or organizations choose open source software are: lower cost, security, no vendor lock in, and better quality.

Artificial neural network has been widely used for solving non-linear complex tasks. With the development of computer technology, machine learning techniques are becoming good choice. The selection of the machine learning technique depends upon the viability for particular application. Most of the non-linear problems have been solved using back propagation based neural network. Errorback propagation based neural network has found applications in various areas. R Tuli [4] applied neural network technique for character recognition and this paper attempts to recognize the characters using a back propagation algorithm. M. M. Subashini

et. al. [5] proposed a neural network technique for classification of brain tumor presence or not. Ngaopitakkul et. al. [6] proposed a decision algorithm for identifying the phase with fault appearance of a two-winding three-phase transformer. A decision algorithm based on a combination of Discrete Wavelet Transforms and backpropagation neural networks (BPNN) is developed. S. A. Radzi et. al. [7] applied 5 layers CNN for recognizing Malaysian license plate. Gradient based back propagation method is used for learning. Choon-Boon Tay et al. [8] proposed a 5 layer CNN for gender classification of pedestrians.

J. Jiang et. al. [9] proposed application of bpnn for pattern recognition purposes. The Back Propagation neural network is used improve the high resolution remote sensing image classification precision.

N. L. D. Khoa et. al. [10] proposed a method to forecast the stock price using neural networks. Both feed forward neural network and simple recurrent neural network, trained by time and profit based back propagation algorithm.

V. Jaiganesh et. al. [11] proposed a novel intrusion detection system (IDS) by back propagation neural networks (BPN). K. U. Rani [12] proposed a neural network technique for classification of heart disease. The neural network is trained with Heart Diseases database by using feed forward neural network model and backpropagation learning algorithm with momentum and variable learning rate. T. A. Anjit et. al. [13] presents a new approach for the detection of nasal bone for ultrasonogram of foetus of 11 to 13 weeks of gestation. The proposed method is based on the extraction of image texture parameter of nasal bone region of ultrasonogram and their subsequent classification using Back Propagation Neural Network (BPNN). From the above, we can draw an inference that Open Source Software is the common demand of today's scenario. It costs zero. Artificial neural network techniques are widely used in solving complex non-linear tasks. Breast cancer detection, character recognition, athletics problems, gender classification, filter designing fault diagnosis of analog circuits and stock price forecasting are some of the wide range of applications employing backpropagation neural network successfully.

Architecture	Best validation performance and @epoch	Training R	Validation R	Testing R	Overall R	Overfitting Problem
4-8-1	0.00281@7	.91679	.92286	.90278	.91548	Yes
4-10-1	0.0030834@56	0.93825	0.93393	0.9496	0.93877	no
4-12-1	0.0047082@4	.90283	.89624	.90096	.90196	yes
4-15-1	.0038416@6	.92826	.91207	.91473	.92278	yes
4-6-4-1	0.0046136@6	.92101	.89448	.92235	.91745	yes
4-8-4-1	0.0031709@8	0.91836	0.92961	0.92013	0.92011	yes
4-7-3-1	0.0049219@13	0.92112	0.89642	0.90633	0.91342	no
4-8-5-1	0.012375@6	0.92529	0.91677	0.91964	0.92303	no
4-10-5-1	0.012123@10	0.92652	0.90658	0.91957	0.92283	yes
4-15-5-1	0.015592@9	0.93049	0.90431	0.92371	0.92471	yes

II. Data Collection and data preprocessing

The data of five open source software is being collected. Table 1 provides the information of this dataset. Type field is the type of software for which it is user for to either convert or to play. File Format is the different files for which the software is supporting and operating system is the system software platform on which this software can run.

Table 1 Data set of five open source softwares

S. No.	Name	Type	Purpose	File Format	Operating System
1	Helium Audio Converter	Audio	Converter	MP3, WMA, WAV, WAVpack, FLAC, OggVorbis, MP4, M4a, AAC, and MPC.	Windows XP/2003/Vista/Server 2008/7/8
2	Fre:AC	Audio	Converter	MP3, MP4/M4A, WMA, OggVorbis, FLAC, AAC, WAV	Windows NT/2000/XP/2003/Vista/Server 2008/7/8
3	DJ Audio Editor	Audio	Editor	MP3, WAV, OGG, WMA, M4A, AAC, AMR, MP2, RA, AC3, FLAC, MKA, AU, AIFF, M4B, M4R	Windows XP/Vista/7/8
4	VLC Media Player	Video	Player	MPEG-1, MPEG-2, MPEG-4, DivX, MP3, OGG	Windows 7, Windows 2003, Windows 8, Windows Vista, Windows XP, Windows Server 2008
5	UM Player	Video	Player	AAC, AC3, ASF, AVI, DIVX, FLV, H.263, Matroska, MOV, MP3, MP4, MPEG, OGG, QT, RealMedia, VOB, Vorbis, WAV, WMA, WMV, XVID	Windows NT/98/Me/2000/XP/2003/Vista/Server2008/7

III. Data Preprocessing

The data has been mapped into numerical form. The 5 software as shown in table 1 under the field name are taken as target class. The features on which selection of this software depends are type, file format, purpose and operating system as given in table 1. The data has been coded in the range of 0 to

9. Also the coded data is normalized before feeding to neural system.

IV. Simulation , Results and discussion

MATLAB NN tool box is used for neural network architecture design. All the five categories of data are fed simultaneously and in random fashion.

Table 2 Various Neural Architecture Results

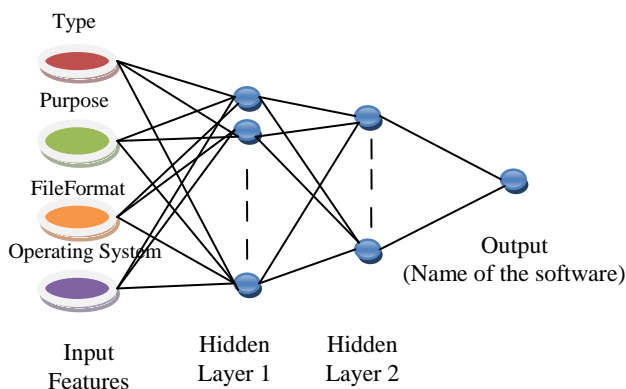


Figure 1 Neural Network Architecture to be designed

Multilayer neural network is designed with single hidden layer and two hidden layer. Figure 1 shows the general architecture of neural system having input features Type, Purpose, Fileformat, Operating system and hidden layers which varied from one to two and output which gives the name of the software. Table 2 provides the different neural architecture results. Architecture is the number of input nodes-number of hidden layer neurons-output.the last value in architecture column given output and in between values gives hidden layer neurons. Best validation performance is the minima obtained by validation dataset. Training, Testing, Validation and overall R is the regression coefficient values for respective dataset. Overfitting problem is measured from the convergence plot. If validation dataset error started increasing after the system achieves best performance, this means the overfitting occurs. This problem should not be for a well trained system. Levenberg-Marquardt training is selected to update the weights. Data is divided in random fashion. The 60% of the dataset is used as training, 20% as testing and 20% for validation. The objective function is mean square error. The μ is kept constant at 0.001. The stopping criterion is 50 continuous iteration checks for validation data, once it has got minima.

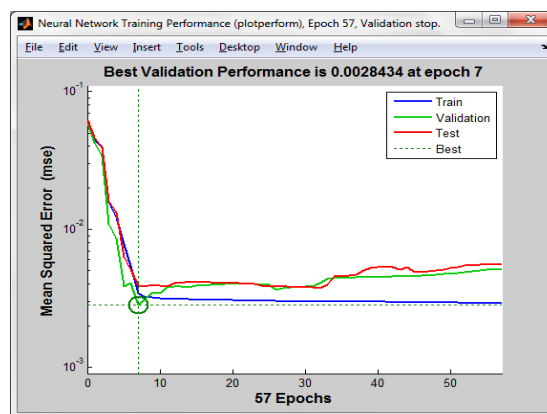


Figure 2 Convergence Plot for 4-8-1

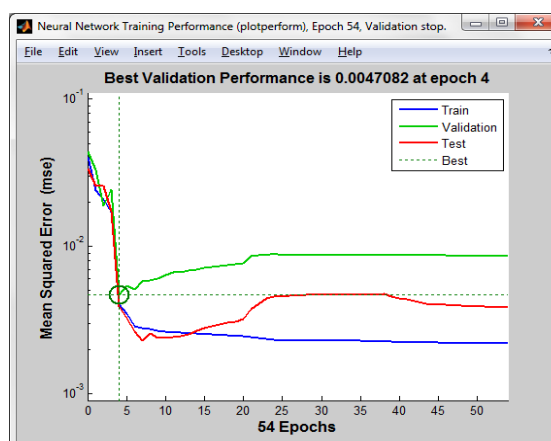


Figure 3 Convergence Plot for 4-12-1

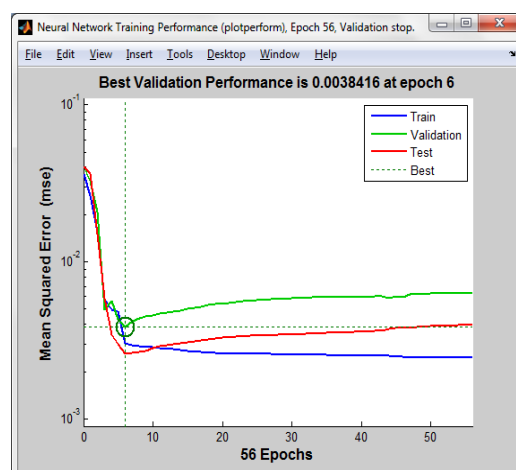


Figure 4 Convergence Plot for 4-15-1

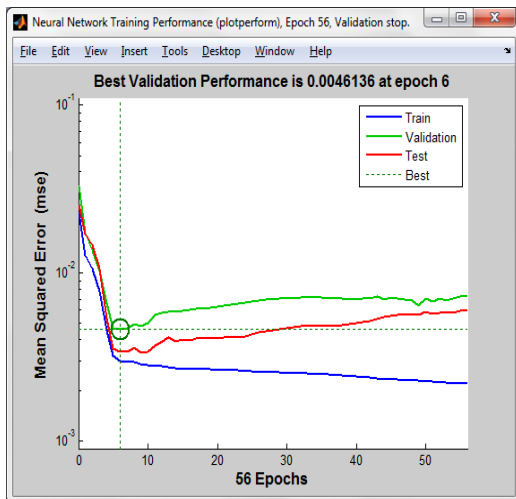


Figure 5 Convergence Plot for 4-6-4-1

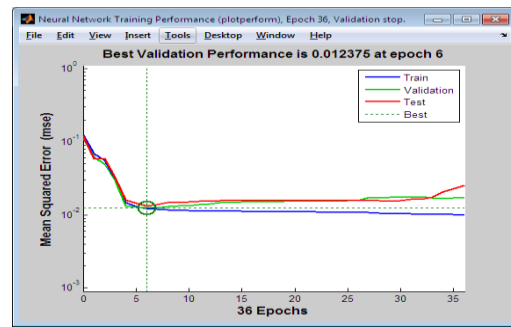


Figure 8 Convergence Plot for 4-8-5-1

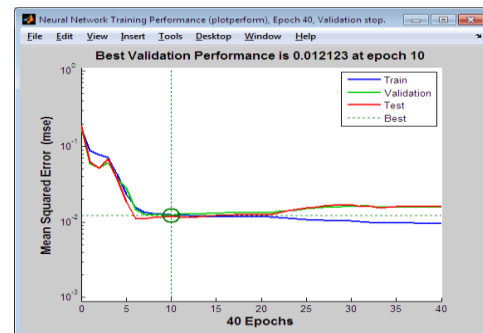


Figure 9 Convergence Plot for 4-10-5-1

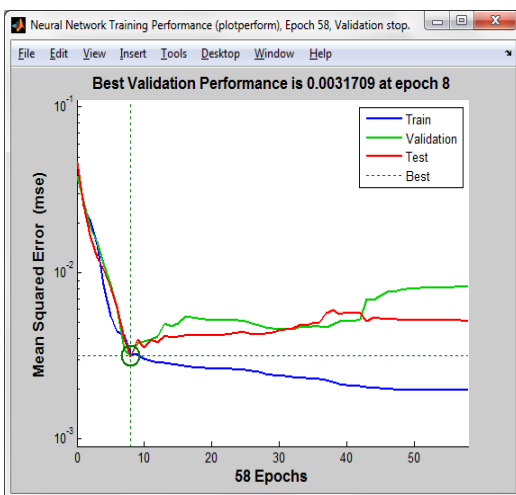


Figure 6 Convergence Plot for 4-8-4-1

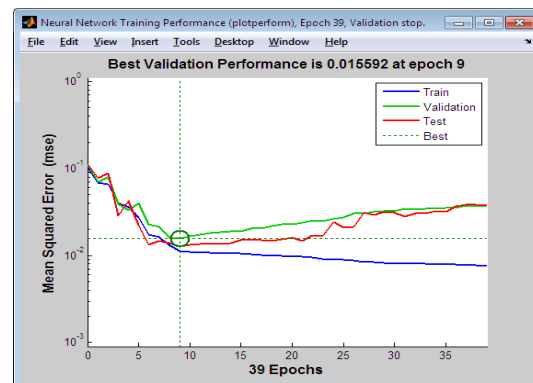


Figure 10 Convergence Plot for 4-15-5-1

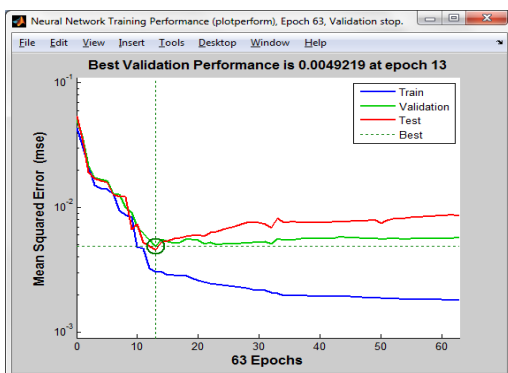


Figure 7 Convergence Plot for 4-7-3-1

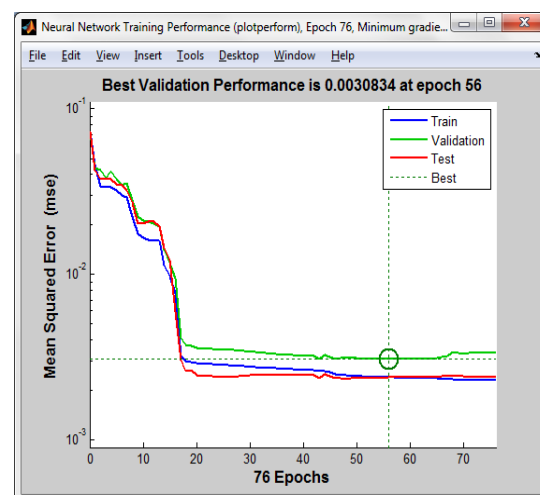


Figure 11 Convergence Plot for 4-10-1

Figure 2 to Figure 11 provides the convergence graph for different architecture. The red highlighted portion in Table 2, the neural system for which the validation error is remaining constant after the best validation performance. For other architecture, the error increases which mean those architectures are having overfitting problem. 4-10-1 has maximum R value than 4-7-3-1 and 4-8-5-1. Also 4-10-1 has minima validation error however this architecture takes more time to achieve minima. Figure 12 shows the regression plot for 4-10-1 architecture.

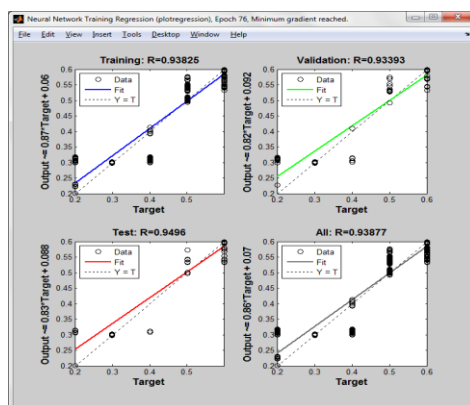


Figure 12 Regression Plot for 4-10-1

V. Conclusion

In this research, the main focus is to make a user friendly environment for online open source software selection. The software selection according to the computer architecture or specification is the problem for users. The installation of unwanted software consumes time. To overcome this problem an automatic system is designed which provides the software according to computer specification. Open source software data is used for automatic system design.

A data of various open source software has been collected and coded into numeric form. Backpropagation based neural network is used as an automatic classification system. Input features from user are accepted by the NN system. The NN system is trained with training dataset. The performance parameter is mean square error. Once the system gets trained, the regression plot gives the measure how efficient the system is. Multilayer network with one hidden layer is providing satisfactory results with fit value of 0.93877.

References

- [1] N. Networks, "Biological Neurons and Neural Networks", *Artificial Neurons*, Neural Networks, pp. 1–14, 2004.
- [2] A. M. St. Laurent, "Understanding Open Source and Free Software Licensing," in *Understanding Open Source and Free*

Software Licensing, First., S. St. Laurent, Ed. O'Reilly Media, 2004, p. 4.

- [3] Irina Guseva, "Bad Economy Is Good for Open Source," 2012. [Online]. Available: <http://www.cmswire.com/>.
- [4] R. Tuli, "Character recognition in neural networks using back propagation method," in *Advance Computing Conference (IACC), 2013 IEEE 3rd International*, 2013, pp. 593–599.
- [5] M. M. Subashini and S. K. Sahoo, "Brain tumour detection using Pulse coupled neural network (PCNN) and back propagation network," *Sustainable Energy and Intelligent Systems (SEISCON 2012), IET Chennai 3rd International on*, pp. 1–6, 2012.
- [6] A. Ngaopitakkul, C. Pothisarn, S. Bunjongjit, J. Klomjit, and B. Suechoey, "Application of back-propagation neural network for transformer differential protection schemes part 2 identification the phase with fault appearance in power transformer," in *Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS), 2012 Joint 6th International Conference on*, 2012, pp. 1499–1503.
- [7] S. A. Radzi and M. Khalil-Hani, "Character recognition of license plate number using convolutional neural network," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7066 LNCS, pp. 45–55, 2011.
- [8] C.-B. Ng, Y.-H. Tay, and B.-M. Goi, "A convolutional neural network for pedestrian gender recognition," in *Advances in Neural Networks-ISNN, Lecture Notes in Computer Science, volume 7951*, Springer International Publishing, 2013, pp. 558–564.
- [9] J. Jiang, J. Zhang, G. Yang, D. Zhang, and L. Zhang, "Application of back propagation neural network in the classification of high resolution remote sensing image: Take remote sensing image of beijing for instance," in *Geoinformatics, 2010 18th International Conference on, IEEE, 2010*, pp. 1–6.
- [10] N. L. D. Khoa, K. Sakakibara, and I. Nishikawa, "Stock price forecasting using back propagation neural networks with time and profit based adjusted weight factors," in *SICE-ICASE, 2006. International Joint Conference, IEEE, 2006*, pp. 5484–5488.
- [11] V. Jaiganesh, P. Sumathi, and S. Mangayarkarasi, "An Analysis of Intrusion Detection System using back propagation neural network," in *Information Communication and Embedded Systems*

- (ICICES), 2013 International Conference on, IEEE, 2013, pp. 232–236.
- [12] K. U. Rani, “Analysis of heart diseases dataset using neural network approach,” in *journal arXiv preprint arXiv:1110.2626*, 2011.
- [13] T. A. Anjit and S. Rishidas, “Identification of nasal bone for the early detection of down syndrome using Back Propagation Neural Network,” in *Communications and Signal Processing (ICCSP), 2011 International Conference on*, 2011, pp. 136–140.