International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5, June 2019

Effectual Predicting Telecom Customer Churn using Deep Neural Network

Bhawna Nigam, Himanshu Dugar, Niranjanamurthy M

Abstract: ***Telecom industry has seen a phenomenal growth throughout the world in recent times. Today companies in this sector are putting their best efforts to retain their churning customers by satisfying them with offers and discounts. It is due to the fact that, acquiring a new customer is far more expensive than retaining an existing one. Deep neural network learns on its own in a supervised manner and thus can be used in this regard efficiently. In this paper we have used the H2o package of deep learning to predict telecom customer churn. H2o package stem from a multi-layer artificial neural network. Number of hidden layers, epoch, number of neurons, hidden dropout ratio, input dropout ratio and activation function have been varied to achieve high sensitivity value. Sensitivity is the percentage of churners who are correctly predicted as churning customers. Our model has achieved sensitivity of 85% and thus the results are satisfactory.***

Index Terms: Telecom Churn Prediction, Deep Learning, Deep Neural Network, H2o platform, Customer churn.

I. INTRODUCTION

Telecommunication industry has turned out to be one of the fastest growing industries and has scaled dramatically in the recent times. Due to the competitive rates of various providers, customers often tend to switch between them and thus calling for the need to predict the potential churners and take appropriate actions to prevent their churning. As a result many attempts have been made in telecom industry to predict the churning customers before they actually leave a provider. In this paper we have implemented deep neural network and then tuned various model parameters to get an optimal model for telecom churn prediction. Various techniques have been used in past for predicting customer churn like machine learning algorithms including Decision trees, Random Forest, Logistic Regression, Support Vector Machine(SVM), Neural Network and many more. Another novel technique to predict telecom customer churn is by using deep neural network. By using it we can build model that fits our data using various hierarchies of concepts thus increasing the performance of the model built. We have used the H2o package available in R for implementing the deep neural network using the K fold cross-validation method. Then we have varied different parameters of the model to achieve high sensitivity value. Sensitivity in this regard is the proportion of actual churners that are correctly predicted by the model.

The term machine learning was first given by Arthur L Samuel[1].Machine learning is subset of artificial intelligence which enables computers to learn (i.e. improvise) from data without any intervention. Machine learning is often coined with the terms pattern recognition and computational learning theory. Machine learning involves constructing algorithms that can learn from data available and can be used to make predictions on data. Here a model is built from the given input data which is then used to make predictions on new data. As the data being collected is drastically increasing each day, this calls for the need of machine learning.

Deep learning is a revolutionary technique in the field of machine learning which is introduced to move machine learning closer to which it was intended for: artificial intelligence [2]. Deep learning is a subset of machine learning that capacitate computers to learn from given data, map and interpret the world in form of a hierarchical concepts. In machine learning, the learning process is extracting knowledge from the given data thus there is no need for a human to specify any kind of knowledge [3].Because of the hierarchy of concepts, complex concepts can be learned by building them from simpler ones thus reducing the complexity. Deep learning is making major advances in solving problems that the artificial intelligence community had tried to solve for many years but was still not able to do that. Today deep learning has applications almost everywhere, for example in image recognition[4], speech recognition[5], prediction's like telecom churn prediction[6], sentiment analysis[7], natural language processing[8], fraud detection[9] and many more. In this paper we have used the multi-layer artificial neural network also called deep neural network (DNN) to predict customer churn. Other types of DNN's include convolution neural network (CNN) and recurrent neural network (RNN). CNN's are used in case of image data whereas RNN's are used in case of sequential data like audio, time series. DNN's are computing systems based on structure and functioning of biological neural networks [10].It is composed of a large number of nodes called neurons similar to the neurons in the brain. Each neuron can transmit a signal to another. The connections between neurons are called edges. Edges have weights that adjust during learning of the neural network. The weight of edge decides the participation of a particular neuron in the deep learning model. Typically, neurons can be divided into three layers: the input layer, hidden layer(s) and output layer. The input layer is composed of the data variables given to the model as input, hidden layers forms hierarchy of concepts and the output layer consists of neurons representing the possible output.

Manuscript published on 30 June 2019.
***** Correspondence Author (s)
Dr. Bhawna Nigam, Assistant Professor, Department of Information Technology, Institute of Engineering & Technology, Devi Ahilya University, Indore, India ,
Mr. Himanshu Dugar, Student, Bachelor of Engineering in Computer Science from Institute of Engineering and Technology(IET DAVV)Indore of batch 2015-19. India.
Dr. Niranjanamurthy M, Assistant Professor, Department of Computer Applications, M S Ramaiah Institute of Technology, Bangalore- 560054, Karnataka, India.

Signals travel from the first layer (the input layer), through the hidden layer to the last layer (the output layer), possibly after traversing the layers multiple times.

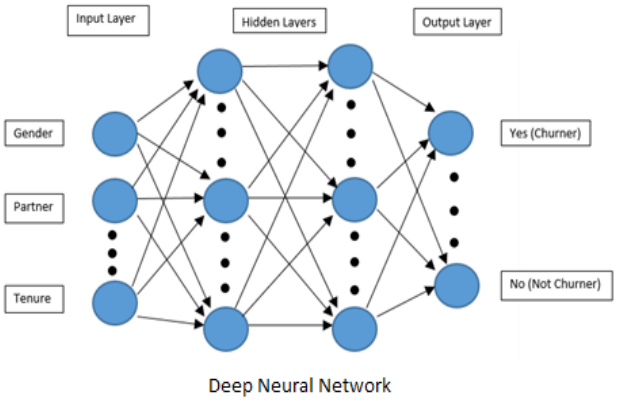

Figure 1: Deep Neural Network

The remaining sections of this paper are organised as follows. Section 2 defines the previous works in the field of telecom churn prediction, Section 3 contains the tools and techniques used in the work done, Section 4 contains the experimental findings and finally section 5 concludes this paper.

II. LITERATURE REVIEW

Prediction of telecom churners has always been an area of researcher's interest and hence many researchers have worked in this direction to predict telecom customer churn.

Hung, Yen and Wang [11] applied data mining techniques like neural network anddecision tree to predict telco churn. They found out that both decision tree and neural network data mining techniques can predict churning customers accurately using various customer details like demographics,billing information, call details and many more. This paper also showed the effect of inadequate data on model building. Due to unavailability of customer demographics Wei and Chiu [12] proposed a technique to identify churners form customers call pattern and information about contracts obtained from customers call details. In this paper multi classifier class combiner approach to solve the problem of skewed distribution of customers as churners and non-churners. Results suggest that the technique provided satisfactory results, as compared to other demographic based churn prediction systems.

T. Vafeiadis, K.I. Diamantaras, G. Sarigiannidis and K.Ch. Chatzisavvas[13] compared the performance of widely used machine learning algorithms including Artificial Neural Networks, Decision Trees, Support Vector Machines, Naïve Bayes classifiers, and Logistic Regression classifiers, then they applied boosting techniques and compared the performance of the boosted versions. The classifier which performed best among the models was the SVM-POLY usingAdaBoost with accuracy of about 97%. Chuanqi Wang, Ruiqi Li, Peng Wang andZonghai Chen [14] discussed that customer churn prediction is cost sensitive problem. In contrast to most of papers which considers each misclassification same, this paper presents that misclassification cost of each sample is different. A partition cost-sensitive CART model is proposed in this paper in order to minimize the cost of churn prediction.

Arno De Caigny, Kristof Coussement and Koen W. De Bock [15] proposed a new composite algorithm built on top oflogistic regressionanddecision treesfor predicting customer churn. They have discussed that decision trees and logistic regression have issues likein handlingvariables that are interacting and many more. The logit leaf model (LLM) is a new algorithm to classify data in an effective manner. The LLM model combines the power of both the base algorithms as first decision tree is used to create segments and then for each leaf model is built. The results that were obtained were better than many advanced methods.

Kiran Dahiya and Surbhi Bhatia [16] proposed a new framework for the customer churn prediction model and which is implemented using WEKA Data Mining software. This paper also compares the performance of decision tree and logistic regression using framework. The results obtained showed that decision tree had much higher accuracy than the logistic regression technique.

Y Huang, F Zhu, M Yuan, K Deng, Y Li, B Ni, W Dai, Q Yang and J Zeng[17] showed how telco big data make prediction churners easier. They have showed that large volume of data, large variety of features and large incoming data improves performance of churn prediction. The proposed system was deployed in one of the biggest mobile operators in china. The system provided prepaid customers who are likely to churn in the next month, having 0.96 precision for the top 50000 predicted churners.

K Kim, C Jun and J Lee [18] proposed a procedure that predicts customer churn by examining communication patterns among customers. The subscribers of telecom companies are connected with other customers, and network properties may affect the churning customers. Hence it makes the use of social network in predicting the churning customers. This paper clearly showed that there was improvement in predicting churners by introducing network analysis as compared to techniques taking personal customer information into consideration's Verbeke, D Martens, C Mues and B Baesens [19] discussed about two advanced data mining techniques which are AntMiner+ and ALBA. They had used these two because: AntMiner+ is high performing technique which allows to incorporate domain knowledge. ALBA brings about high accuracy prediction of a non-linear SVM. The results showed that use of ALBA results in increased performance by improving learning of classification techniques and AntMiner+ resulted in accurate and comprehensible model.

Y Zhang, R Liang, Y Li, Y Zheng and M Berry[20] presented prediction system based on behaviour to predict telecom churn. This system derives attributes from services that the customer avails. Hence only usage pattern is taken into consideration to make prediction, eliminating the problems like feature selection, missing values and many more. Also this technique was free from the problems that traditional systems have to face like correlations among inputs. The proposed system made use a clustering algorithm to predict customer churn.Y Huang and T Kechadi [21]proposed a hybrid model that combines both supervised and unsupervised techniques to predict customer churn.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

This hybrid model combines modified k means clustering algorithm and classic rule inductive technique (FOIL). The experimentation included verifying that weighted k-means clustering can lead to better partitioning, comparison of classification results with other well-known techniques and finally comparison of proposed model with other hybrid classification techniques. Results showed that the proposed model had maximum accuracy over the benchmarking data used.

III. TECHNIQUES AND TOOLS

Deep neural network (DNN)

Although the exact definition of learning is difficult to formulate, but in context to DNNor feed forward ANN, it can be seen as updating the network architecture i.e. the links and weights to perform a particular task. The network learns the weights from the training examples given to it and update them iteratively till the convergence criterion is met [22].Hence the deep neural network learns by itself rather than following rules specified by a human.

H2o package

In this paper we have used h2o package in R to build the artificial neural network.H2o platform has many benefits which are not restricted to it being open source. It also includes advantages like it is fast, scalable and H2o is easier to use as it has interfaces for many languages including R and Python [23]. H2o include many machine learning algorithms like logistic regression, naïve bayes etc and also implements deep learning algorithms.

Some of H2o's deep learning features include [24]-

1. supervised training protocol for regression and classification modelling tasks
2. As h2o's core code is in java so it provides fast and memory efficient Java implementations as data is stored in compressed way.
3. Automatic and adaptive learning rates for fast convergence
4. Regularization options such as l1, l2 and dropout to prevent over-fitting
5. completely scriptable R API from H2O's CRAN package

Cross validation

In the process of finding an optimal model to predict customer churn in telecom industry we require 3 datasets which are as follows: 1. Training set - The training set which consists of input examples is used to build the model. This set consists of examples used to fit the parameters (i.e. the weights of links between neurons in the artificial neural network).

2. Validation Set /Dev Set- This dataset is used to tune the model by varying various parameters to improve the model's performance. It is kept different from training set because if we use the training set to tune the model so it will actually learn the training set. So it would make predictions with great accuracy on training data but will have very poor performance on data which it has not seen before (i.e. other than training data). So it have poor generalization and hence of no use as in real life applications we require the model to make predictions on data which it has not seen before.

3. Test Set- This dataset is used to test the actual performance of model. This consists of data which is not used in either training or tuning the model.

Here we have used K fold Cross Validation method which is as follows: It is used to validate a model without having an explicit validation set. Consider k=5. So in total 6 models are built. First 5 are built on 80% training data and 20% is held out for each model. It can be depicted as follows:

Model 1

V1	Tr	Tr	Tr	Tr

Model 2

Tr	V2	Tr	Tr	Tr

Model 3

Tr	Tr	V3	Tr	Tr

Model 4

Tr	Tr	Tr	V4	Tr

Model 5

Tr	Tr	Tr	Tr	V5

Where Tr-Training Set, V-Validation Set (or Holdout Set)

Finally the main model is built as follows: The 5 holdout predictions are combined into single prediction and then this is compared with the training set i.e. holdout predictions are compared with true training labels, but the model which predicted that holdout set had not seen it in its training set.

True Labels	True Labels	True Labels	True Labels	True Labels

Holdout Set(V1)	Holdout Set(V2)	Holdout Set(V3)	Holdout Set(V4)	Holdout Set(V5)

By comparing the above two sets we get the overall cross validation matrix.

Regularization

Regularization is a technique which is used to prevent over-fitting of model on training data. Here we have used l1 regularization (l1 stands for Lasso) has the effect of reducing weights (to zero) to prevent over-fitting. Another advanced technique for regularization is dropout. We have used hidden dropout ratio and input dropout ratio parameters as well in the model to improve generalization.

IV. EXPERIMENTS AND RESULT ANALYSIS

For the experiment we have taken dataset containing 30 predictor variables that are: gender, Senior Citizen, Partner, Dependents, tenure, Phone Service, Paperless Billing, Monthly Charges, Total Charges, Multiple Lines No phone service, Multiple Lines Yes, Internet Service Fiber optic, Internet Service No, Online Security No internet service, Online Security Yes, Online Backup No internet service, Online Backup Yes, Device Protection No internet service, Device Protection Yes, Tech Support No internet service, Tech Support Yes, Streaming TV No internet service, Streaming TV Yes, Streaming Movies No internet service, Streaming Movies Yes, Contract One year, Contract Two year, Payment Method Credit card (automatic), Payment Method Electronic check, Payment Method Mailed check and 1 variable to be predicted i.e.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

Churn (having 2 possible values: yes or no).In total the dataset has 4914 rows which is divided into training set and test set as 70% and 30%.In the experiment we have chosen a base model.

The base model has the following parameters and values: x: Churn, y: remaining 30 predictor variables, churnTrain: Training dataset.

```
model<- h2o.deeplearning(x=x,
                         y=y,
training_frame = churnTrain,
seed=123,
variable_importances = TRUE,
activation="MaxoutWithDropout",
hidden = c(90),
hidden_dropout_ratio = c(0.1),
                         l1 = 1e-5,
nfolds=6,
initial_weight_distribution = "Normal",
balance_classes = TRUE,
sparse=TRUE,
reproducible = TRUE,
epochs = 10,
keep_cross_validation_predictions = TRUE)
```

The h2o.deeplearning can take a large number of parameters. For our research work, we have selected some of them and we vary these parameters of h2o.deeplearning to tune the artificial neural network. At each stage we choose one model as the base model for further tuning (highlighted).

1. **Number of neurons**

   It is seen in most of the literatures that the number of neurons in the neural network should be 3 times the number of input variables. Hence in the original model there are 90 neurons. Now we change the number of neurons and observe its effects.

Model		Training Set		Dev Set				Variance (Dev err-Train err)
	No. of neurons	Bias (train err.)	AUC	Dev Err.	AUC	Accuracy	Sensitivity	
1	***90***	***24.14***	***82.60***	***28.97***	***77.22***	***71.03***	***78.98***	***4.83***
2	50	23.24	84.26	25.96	80.30	74.04	72.37	2.72
3	30	23.18	85.18	25.39	81.55	74.61	74.85	2.21
4	150	22.19	82.05	25.54	78.66	74.46	71.31	3.35
5	200	22.33	80.95	27.31	78.06	72.69	77.44	4.98

Figure 2: Number of Neurons

Model 1 is selected as base model for further tuning.

As our aim is to predict the customer churn correctly hence we need the sensitivity to be high. The confusion matrix can is as follows:

Actual / Predicted	Churn customer	Not churn customer
Churn customer	True positive	False negative
Not churn customer	False positive	True negative

Figure 3: Actual and predicted Churn customer and Non Churn customer

Sensitivity can be defined as:

Sensitivity=True Positive/Positive, where Positive=True Positive+ False Negative

So we get the highest sensitivity in model 1 (having number of neurons= 90) i.e. 78.98, hence we select it as base model for the future tuning work. Next we observe results by changing the number of hidden layers present in the artificial neural network.

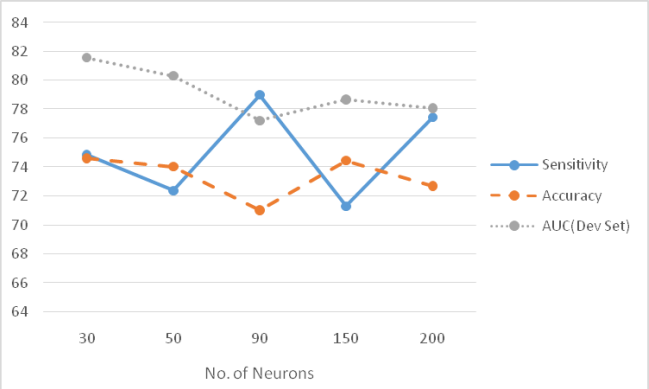

Figure 4:Analysis of umber of Neurons

2. **Number of layers**

Model	No. of layers	Layers	Training Set		Dev Set				Variance (Dev err-Train err)
			Bias (train err.)	AUC	Dev Err.	AUC	Accuracy	Sensitivity	
1	***1***	***c(90)***	***24.14***	***82.60***	***28.97***	***77.22***	***71.03***	***78.98***	***4.83***
2	2	c(90,90)	21.17	81.69	23.62	77.17	76.38	70.12	2.54
3	2	c(50,50)	23.56	80.96	27.79	76.37	72.21	78.86	4.23
4	3	c(90,90,90)	21.81	80.28	21.83	72.77	78.17	60.33	0.02
5	3	c(50,50,50)	22.45	79.70	24.22	76.42	75.78	67.53	1.77

Figure 5:Number of Layers

Model 1 is selected as base model for further tuning.

The observations recorded in the above table can be seen as: Both model 1 and 3 have almost same sensitivity (model 2, 4, 5 have low sensitivity values), we select model 1 as base model for further tuning because it has good combination of Area under the curve values for both training as well as test sets, accuracy and sensitivity.

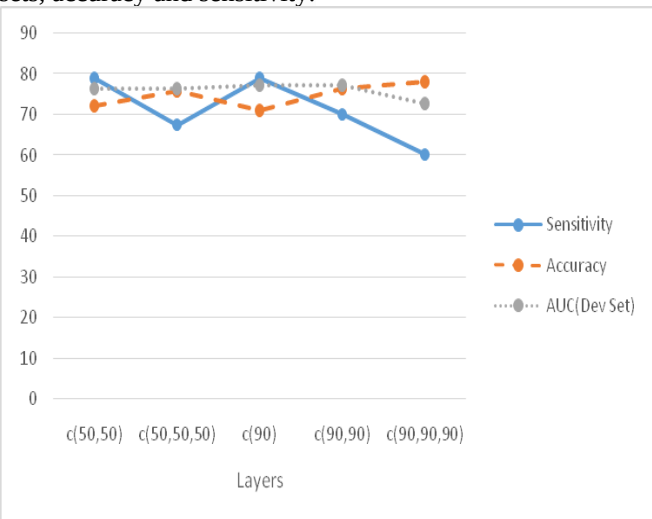

Figure 6: Analysis of Layers

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

3. Epoch

The bias i.e. training error is high and we need to tackle it and so we increase the number of epoch. Epoch is the number of times model goes through the input examples given to it and thus resulting in reduced training error.

Model	Epoch	Training Set		Dev Set				Variance (Dev err-Train err)
		Bias (train err.)	AUC	Dev Err.	AUC	Accuracy	Sensitivity	
1	10	24.14	82.60	28.97	77.22	71.03	78.98	4.83
2	20	22.95	84.75	29.90	79.41	70.10	80.51	6.95
3	25	22.57	86.40	28.76	80.55	71.24	78.98	6.19
4	50	20.74	87.77	21.79	83.53	78.21	69.30	1.05
5	100	20.74	87.77	21.79	83.53	78.21	69.30	1.05

Figure 7:Epoch

Model 2 is selected as base model for further tuning.

The observations of above table can be interpreted as: Model 4 and 5 have low sensitivity values. In case of model 1, 2 and 3, we choose model 2 for further tuning because it has good combination of sensitivity (80.51 i.e. our highest priority) and bias (lower than model 1 and a bit higher than model 3). But as it can be seen that with decrease in bias, the variance has increased. So there is trade-off between variance and bias.

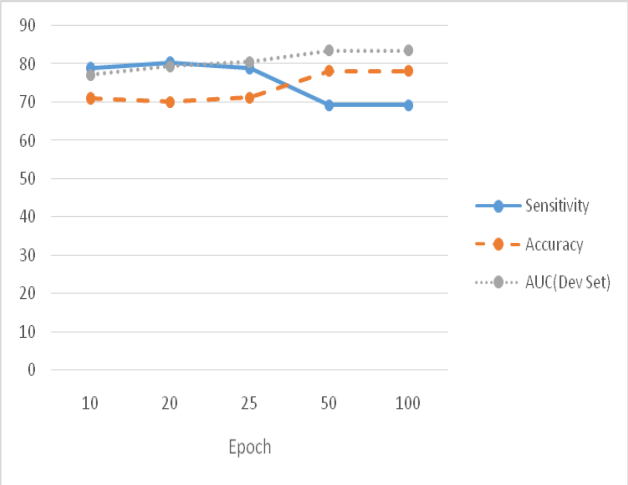

Figure 8: Analysis of Epoch

4. Hidden dropout ratio

Now we apply regularisation by using different values for hidden dropout ratio. The hidden dropout ratio defines the fraction of inputs to omit by every hidden layer to improve generalization i.e. increasing the ability of model to make predictions correctly on data which it has never seen before.

Model	Hidden dropout ratio	Training Set		Dev Set				Varia (De err-T ert
		Bias (train err.)	AUC	Dev Err.	AUC	Accuracy	Sensitivity	
1	0.09	22.07	85.18	29.21	79.54	70.79	79.22	7.1
2	0.10	22.95	84.75	29.90	79.41	70.10	80.51	6.9
3	0.11	22.93	84.65	27.94	79.83	72.06	76.74	5.0
4	0.15	23.64	84.55	26.23	80.47	73.77	75.32	2.59
5	0.20	23.46	85.11	26.35	80.81	73.65	76.03	2.89

Figure 9: Hidden dropout ratio

Model 2 is selected as base model for further tuning.

Here we see that model 2 has best combination of sensitivity value and variance as models 3, 4 and5 have less sensitivity value and model 1 has high variance. Hence we select model 2 for further tuning.

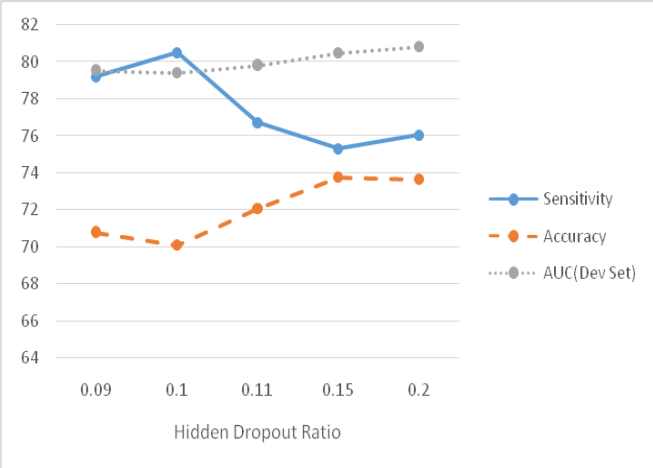

Figure 10: Analysis of Hidden dropout ratio

1. Input dropout ratio

We also use Input dropout ratio to enhance regularization. The input dropout ratio defines the fraction of feature for each row of training examples to be omitted from training to improve generalization.

Model	Input dropout ratio	Training Set		Dev Set				Variance (Dev err-Train err)
		Bias (train err.)	AUC	Dev Err.	AUC	Accuracy	Sensitivity	
1	0.01	23.30	84.34	29.75	79.72	70.25	80.75	6.45
2	0.02	23.58	84.51	26.59	80.04	73.41	74.73	3.01
3	0.03	24.16	84.03	26.02	80.44	73.98	72.96	1.86
4	0.04	23.26	84.09	24.88	80.64	75.12	72.55	1.62
5	0.05	23.60	84.23	24.49	80.63	75.51	70.12	0.89

Figure 11: Input dropout ratio

Model 1 is selected as base model for further tuning.

The above table clearly shows that that model 1 has highest sensitivity value i.e. 80.75 and rest of the models are having sensitivity in the range 70 to 75, hence selecting model 1 for further tuning.

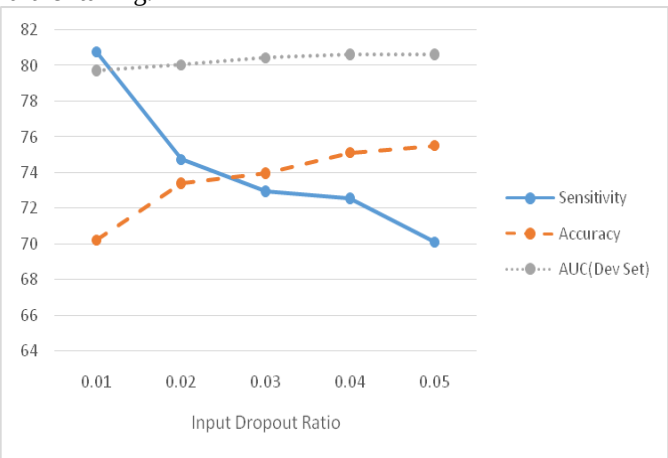

Figure 12: Analysis of Input dropout ratio

1. Activation Function

Here we use different activation functions and observe the results.

M		Training Set	Dev Set	Varianc

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

o d e l	Activation Function	Bias (train err.)	AUC	Dev Err.	AU C	Accur acy	Sensit ivity	e (Dev err-Tra in err)
1	MaxoutWithDropout	23.30	84.34	29.75	79.72	70.25	80.75	6.45
2	TanhWith Dropout	22.67	85.94	24.76	82.44	75.24	74.26	2.09
3	RectifierWithDropout	26.79	79.75	28.27	76.78	71.73	72.49	1.48

Figure 13: Activation Function

As it is clearly visible that use of Maxout With Dropout gives highest sensitivity value i.e. 80.75, hence we select model 1.

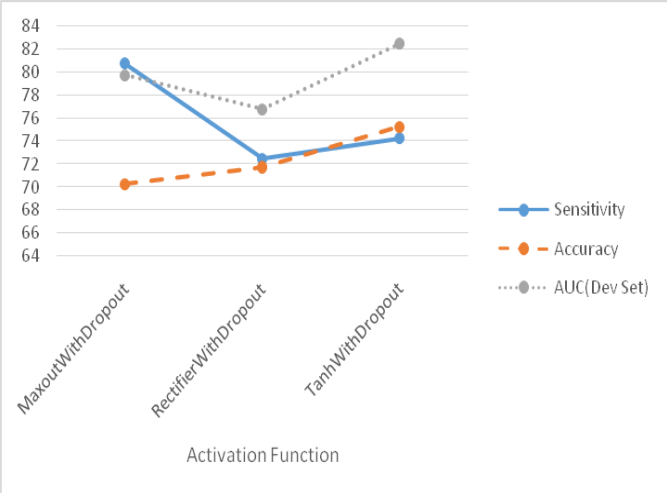

Figure 14: Analysis of Activation Function

Now this model is used to predict the values for churn on the test set, which the model has not seen till now. The confusion matrix containing results of predictions on the test data is as follows.

Prediction results on test set:

Prediction / Actual	No	Yes	Total
No	746	66	812
Yes	409	362	771
Total	1155	428	1583

Total error= (66+409)/1583=30.01%
Accuracy=69.99%
Sensitivity= 84.58%
Specificity= 64.59%

V. CONCLUSION

Hence we have obtained the optimal model for prediction of telecom churn by varying the parameters. It is taken care that the model does not over fit the training data and also has a good sensitivity value i.e. 84.58% (i.e. model is generalized) on the test dataset. Hence we have obtained satisfactory results as compared to traditional methods for predicting customer churn. Telecom industry has seen a phenomenal growth throughout the world in recent times. Today companies in this sector are putting their best efforts to retain their churning customers by satisfying them with offers and discounts. It is due to the fact that, acquiring a new customer is far more expensive than retaining an existing one. Deep neural network learns on its own in a supervised manner and thus can be used in this regard efficiently. In this paper we have used the H2o package of deep learning to predict telecom customer churn. H2o package stem from a multi-layer artificial neural network. Number of hidden layers, epoch, number of neurons, hidden dropout ratio, input dropout ratio and activation function have been varied to achieve high sensitivity value. Sensitivity is the percentage of churners who are correctly predicted as churning customers. Our model has achieved sensitivity of 85% and thus the results are satisfactory.

REFERENCES

1. Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3), 210-229. 1959
2. KB, S. K., Krishna, G., Bhalaji, N., &Chithra, S. (2018). BCI cinematics–A pre-release analyser for movies using H2O deep learning platform. Computers & Electrical Engineering. 2018
3. Goodfellow, I., Bengio, Y., Courville, A., &Bengio, Y. (2016). Deep learning (Vol. 1). Cambridge: MIT press. 2016
4. Tompson, J. J., Jain, A., LeCun, Y., &Bregler, C. (2014). Joint training of a convolutional network and a graphical model for human pose estimation. In Advances in neural information processing systems (pp. 1799-1807).2014
5. Zhang, Y., Chan, W., &Jaitly, N. (2017, March). Very deep convolutional networks for end-to-end speech recognition. In Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on (pp. 4845-4849). IEEE. March 2017
6. Prashanth, R., Deepak, K., &Meher, A. K. (2017, July). High accuracy predictive modelling for customer churn prediction in telecom industry. In International Conference on Machine Learning and Data Mining in Pattern Recognition (pp. 391-402). Springer, Cham. July 2017
7. Araque, O., Corcuera-Platas, I., Sanchez-Rada, J. F., & Iglesias, C. A. (2017). Enhancing deep learning sentiment analysis with ensemble techniques in social applications. Expert Systems with Applications, 77, 236-246. 2017
8. Young, T., Hazarika, D., Poria, S., & Cambria, E. (2017). Recent trends in deep learning based natural language processing. arXiv preprint arXiv:1708.02709. 2017
9. Wang, Y., & Xu, W. (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. Decision Support Systems, 105, 87-95. 2018
10. Van Gerven, M., &Bohte, S. (Eds.). (2018). Artificial neural networks as models of neural information processing. Frontiers Media SA. 2018
11. Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. Expert Systems with Applications, 31(3), 515-524. 2006
12. Wei, C. P., & Chiu, I. T. (2002). Turning telecommunications call details to churn prediction: a data mining approach. Expert systems with applications, 23(2), 103-112. 2002
13. Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., &Chatzisavvas, K. C. (2015). A comparison of machine learning techniques for customer churn prediction. Simulation Modelling Practice and Theory, 55, 1-9. 2015
14. Wang, C., Li, R., Wang, P., & Chen, Z. (2017, July). Partition cost-sensitive CART based on customer value for Telecom customer churn prediction. In Control Conference (CCC), 2017 36th Chinese (pp. 5680-5684). IEEE. July 2017
15. De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. European Journal of Operational Research, 269(2), 760-772. 2018
16. Dahiya, K., & Bhatia, S. (2015, September). Customer churn analysis in telecom industry. In Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), 2015 4th International Conference on (pp. 1-6). IEEE. September 2015
17. Huang, Y., Zhu, F., Yuan, M., Deng, K., Li, Y., Ni, B., ...& Zeng, J. (2015, May). Telco churn prediction with big data. In Proceedings of the 2015 ACM SIGMOD international conference on management of data (pp. 607-618). ACM. May 2015

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

18. Kim, K., Jun, C. H., & Lee, J. (2014). Improved churn prediction in telecommunication industry by analyzing a large network. Expert Systems with Applications, 41(15), 6575-6584. 2014
19. Verbeke, W., Martens, D., Mues, C., &Baesens, B. (2011). Building comprehensible customer churn prediction models with advanced rule induction techniques. Expert Systems with Applications, 38(3), 2354-2364. 2011
20. Zhang, Y., Liang, R., Li, Y., Zheng, Y., & Berry, M. (2011, July). Behavior-based telecommunication churn prediction with neural network approach. In Computer Science and Society (ISCCS), 2011 International Symposium on (pp. 307-310). IEEE. July 2011
21. Huang, Y., &Kechadi, T. (2013). An effective hybrid learning system for telecommunication churn prediction. Expert Systems with Applications, 40(14), 5635-5647.2013
22. Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., &Vandergheynst, P. (2017). Geometric deep learning: going beyond euclidean data. IEEE Signal Processing Magazine, 34(4), 18-42. 2017
23. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/welcome.html (Web link reference)2019
24. Candel, A., Parmar, V., LeDell, E., & Arora, A. (2016). Deep learning with H2O. H2O. ai Inc. 2016

AUTHORS PROFILE

Dr.Bhawna Nigam received Ph.D. in Computer Engineering from Devi Ahilya University, Indore, M.P. in 2017, M.E. in Software Engineering with distinction from Institute of Engineering & Technology (IET), Devi Ahilya University in 2008 and B.E in 2003 from Institute of Engineering & Technology (IET), Devi Ahilya University, Indore. She is currently with Institute of Engineering and Technology (IET), Devi Ahilya University, Indore, India as Assistant Professor in Information Technology department. She is with Devi Ahilya University since 2007. Her area of interest are Machine Learning, Deep Learning, Data Mining, Big Data.She has published 20+ papers on these topics.

Himanshu Dugar pursuing Bachelor of Engineering in Computer Science from Institute of Engineering and Technology(IET DAVV)Indore of batch 2015-19. He has been participating in various Workshops/Conferences on different aspects related to Computer Science. His areas of interest include Data Science, Machine Learning, IoT, Big Data analytics and Cloud Computing. Being passionate about programming, he have won many National level Competitions, Hackathons and is a keen learner having great interest in Research and Development.

Dr. Niranjanamurthy M received Ph.D. Computer Science degree from JJT University, Rajasthan, INDIA in the year 2016, Mhil-Computer Science degree from VM University, Tamil Nadu in the year 2009. MCA degree from Visvesvaraiah Technological University, Karnataka in the year 2007 and BCA Degree from Kuvempu University in the year 2004. He is an Assistant Professor in the department of Computer Applications, M S Ramaiah Institute of Technology, Bangalore. His areas of interests are software testing, e-commerce and m-commerce, software engineering, web technologies, Cloud Computing, Big data analytics, Blockchain Technology, Networking he has been participating in national and international workshops/Conferences on different aspects related to Computer Applications. Guiding Research Scholars, Recognized PhD research examiner of various Universities (National and International) . Published many research Articles related to Computer Science.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)