

Financial Predictions Using Cost Sensitive Neural Networks for Multi-class Learning

Eleni Rozaki

School of Computer Science and Informatics Cardiff University, Cardiff, United Kingdom

e.rozaki@cs.cardiff.ac.uk

Keywords: Multi-class learning, Artificial Network Modelling, Multilayer Perceptron, Radial Basis Function, Cost Sensitive Neural Networks.

Abstract— The interest in the localisation of wireless sensor networks has grown in recent years. A variety of machine-learning methods have been proposed in recent years to improve the optimisation of the complex behaviour of wireless networks. Network administrators have found that traditional classification algorithms may be limited with imbalanced datasets. In fact, the problem of imbalanced data learning has received particular interest. The purpose of this study was to examine design modifications to neural networks in order to address the problem of cost optimisation decisions and financial predictions. The goal was to compare four learning-based techniques using cost-sensitive neural network ensemble for multiclass imbalance data learning. The problem is formulated as a combinatorial cost optimisation in terms of minimising the cost using meta-learning classification rules for Naïve Bayes, J48, Multilayer Perceptions, and Radial Basis Function models. With these models, optimisation faults and cost evaluations for network training are considered.

Introduction

The issue of class imbalance has been identified as a major and important problem in machine learning and data mining. The issue arises when the training data are not distributed evenly between classes. The problem is of even greater concern in real-world applications, especially fraud detection and medical diagnosis. For the purposes of fraud detection and medical diagnosis, standard classifiers have shown very poor performance. The reason for the poor performance is that the classifiers become overwhelmed by the major class, which results in ignoring the minority class. What is demonstrated is that traditional algorithms need to be improved so that they are able to use imbalanced data [1]

The process of calculating and adjusting the cost probability threshold of a network fault with classifiers for meta-cost sensitive learning provides strong precision regarding the minimum costs associated with fault detection and network optimisation. This process can occur based on two dimensions, which are Bayesian threshold for network fault classes and cost classes. In this regard, the model that is proposed in this investigation is based on an evaluation of system performance on causes of KPIs through the use of cost estimation of network issues. This allows for the determination of a set cost matrix for carrying out cost optimization decisions. The cost sensitive classification technique that is used is performed with weka tools, which is an accepted and often-used tool for pattern recognition and classification.

Cost-sensitive solutions allow for a focus on reducing misclassification costs. The problem, however, is that many of the machine learning problems and algorithms that exist are binary. In order to improve how they function and to allow for efficient cost sensitivity analysis for network fault detection, the algorithms need to be changed so that they can be generalised to deal with problems that have many classes as opposed to a single class [2].

The original goal for this investigation was to examine if classification cost was impacted by the cost-sensitive approach, as well as the J48 rules of class. A comparison was performed of the classification cost incurred with a traditional classification of a cost-insensitive approach with four cost-sensitive classifications (CSC) with Bayes, and J48 Multilayer Perceptron (MP) and Radial

Basis Function (RBF). The inclusion of J48 was for a demonstration of the meta-cost consideration for a reduction in costs associated with misclassification [3].

To address that, we stated a set of design criteria that were used to guide the development of the cost optimisation evaluation process. These included firstly various types of KPIs that have different sensitivities to network performance and cost minimisation to support finance decisions. Secondly, the pure cost optimisation studied in this paper that provides a lower bound on the bandwidth cost achievable that simultaneously optimises network performance with the minimum cost of a non-monetised expression of values. Ultimately, the optimisation faults valuation while also enclosing the values of multiple types of equipment needed for various network malfunctions such as transceivers (TRX) for the transmission and reception of several radio frequency (RF) signal issues or antennas replacement for infrastructure costs including in warning status category of faults.

Cost Sensitive and Neural Networks

It is certainly imperative for network administrators and the companies that operate networks to focus on the issue of cost. Such networks are already very expensive to operate, and can become even more expensive if optimisation is not achieved. Network operators and administrators need a way to handle neural networks in a way that reduces misclassification costs and that can allow for imbalanced data to be handled effectively and efficiently.

At the present time, imbalanced training data are generally handled through the process of re-sampling. However, re-sampling has a consequential downside, which is the loss of information and model overfitting as the data distribution are changed. The most efficient and effective manner of handling class imbalance is to assign distinct costs to the training examples. Cost-sensitive learning is able to account for misclassification costs in the model construction. In addition, cost-sensitive learning does not require that imbalanced data be modified directly [1].

The primary goal and target of neural networks is the estimation of a particular requested function. Hidden patterns are classified into classes that are based on prior examples. This creates a discrete output, which is different than with non-parametric regression models in which the outcome is continuous values [4]. The goal of this project was to find the hidden values of key performance indicators and review the abnormal value of network costs. Radial Basis Function-RBF networks of supervised learning are a specific type of linear model, capable to perform regressions, classifications and time series predictions.

Network Performance vs. Cost

Methods for achieving network performance have traditionally been used by network administrators and technicians as a means of finding network faults and returning networks to conditions of optimisation. Optimisation methods have not been widely used by managers as a means of approaching financial predictions. However, the multiclass cost-sensitive classification problem that aims at achieving cost efficient network improvement does require predicting cost optimisation rules with regards to fault detection. Financial performance in the process of broader network fault detection and optimisation is possible [5].

The cost optimisation that is considered as the foundation of this investigation is indeed an important part of the larger process of network performance improvement for the following reasons: First, a cost sensitive learning classifier algorithm can have a general comprehension of cost efficiency based on network performance improvement. Secondly, different KPIs have different sensitivities to both performance and cost. The idea of sensitivities to both cost and performance is important because customers are generally willing to pay more for the use of a network when performance is optimised. Finally, the cost optimisation that is the basis for this investigation is about bandwidth costs in relation to the improvement of network performance and overall cost. In other words, the variance in cost with regards to different types of traffic on a network is vital to the larger cost structure of actual service and content delivery [5].

Cost Allocations and Fault Detection

The rapid changes in technologies combined with the high level of competition that exists in many telecommunications markets creates a situation in which companies in the industry have an incentive to take advantage of their monopolistic advantages as a means of subsidising their activities. Even more, state regulatory agencies that attempt to implement residual revenue requirements are unlikely to achieve their desired outcomes because such revenue requirements are shifted from competitive services to monopolistic services [6].

The creation of a test to determine appropriate and proper costs and to prevent cross-subsidisation is very difficult in the telecommunications industry due to the pervasive common and joint costs. Telecommunications companies generally supply multiple services together, rather than providing individual services that are unrelated to each other. The result is that investment costs are common to multiple services. From a theoretical standpoint, cost allocation must be based on the level to which each service that is provided uses the costs that are incurred [6].

In terms of algorithms, classification algorithms can be used to guide the learning process toward the minority class. This adaptation can occur by changing an algorithm in favour of a minority class, or through the creation of an entirely new algorithm. In addition, the consideration of misclassification costs for each class can also occur for both the algorithms that are used and the data that are used [7].

Cost Benefit Analysis in Multiple KPI Class Prediction

Common costs and directly attributable costs include fixed costs and variable costs. In an optimisation problem, fixed cost only shifts the cost curve into a higher or lower level. So, we only considered variable cost in our analysis. In building and maintaining the sensor system, these expenses correspond to hardware installation and maintenance, data collection, data processing and data analysis. These expenses would include, Traffic cost, Transmission cost, Infrastructure cost, and Maintenance and Operation costs for the optimised cells [8].

C1: Cost of Infrastructure. The cost of infrastructure involves the purchasing of proper equipment for data collection, data processing and data analysis [7]. The KPI used in this investigation to evaluate infrastructure cost is the Handover Success Rate indicator (HSR). The HSR is a measure of the interference, either internal or external, which may effect call switching [9].

HSR degradation can occur as a result of several issues and problems. The degradation can occur due to missing adjacencies, as well as due to hardware faults. Furthermore, location area code boundaries may be incorrectly planned or defined. Finally, coverage limitations can also result in HSR degradation. In order to improve HSR, a spectrum analysis may be performed to find a band with reduced interference. Adjacencies audits might also be performed, as well as improvements to coverage areas. Parameter modifications to handover margins, traffic handover, and power budgets might also occur as a means of improving HSR degradation [9].

C2: Cost of Traffic. The cost of traffic is the expenses associated with installing sensor stations which are located in geographically dispersed areas [8]. The model investigated in this paper includes the traffic issued through an examination of fault cells effected by the traffic channel congestion rate (TCHCR) indicator. The TCHCR is an important KPI for achieving network optimisation because it is an indicator of the rate of congestion across traffic channels [9].

Traffic channel congestion can occur due to hardware faults, as well as due to increases in the number of subscribers and increase in the amount of traffic on a network. Reducing the TCHCR can be achieved by increasing the number of transceivers and transmission equipment on a network, as well as replacing or repairing faulty hardware. In addition, companies can use movable equipment as a means of reducing congestion in areas in which it is predicted that network traffic will increase, such as near as sports venue during a major sporting event [9].

C3: Cost of Transmission. The cost of transmission is the cost of maintaining a system or network. These costs include both regular maintenance of the equipment and periodic replacement of equipment [8]. The measure of the cost of transmission in relation to network performance is the call set-up success rate (CSSR), which is the rate of success of calls being established in comparison to the failure rate of calls being established, and call drop rate (CDR), which is the ratio of calls that are completed successfully as compared to calls that are dropped while in progress [9].

The CSSR may be impacted by radio interface congestion, a lack of radio resources, an increase in traffic, and faulty hardware. In addition, CSSR can increase when there are transmission limitations. In order to diagnosis an increase in CDR, radio uplink statistics may be monitored, along with path balance statistics, which are indicators and measures of uplink and downlink power on the network. Customer complaints regarding blocked calls should also be reviewed as an indicator of CDR problems. Furthermore, spectrum scanners can also be helpful in finding where CDR problems exist, as well as drive test reports. Once these measures are obtained and examined, the way in which to improve CDR problems is to replace faulty hardware [9].

C4: Cost of Maintenance system. Maintenance includes both regular and periodic maintenances such as replacements of necessary equipment and upgrades [10]. Computers maintenance and software upgrades, from sensors to a central computer, including the cost of laying out the wired or wireless network and its own maintenance costs. We estimated the cost of maintenance equipment and operations with the evaluation of the cells that we consider as “norm” in previous research [12] so cells that are optimised but still have a cost estimation regarding the salaries, operation and maintenance of equipment.

Artificial Neural Network Predictions for Optimising Decisions

The network architecture of artificial networks can be identified as either feed forward or recurrent type. The four cost classes of neural networks can be structured in different configurations. A multilayer perception (MLP) network is a network in which the output of one layer is connected to the input of the next layer. This connection occurs through a synaptic weight. The recurrent type can have at least one feedback connection between the neurons that are in the same layer or in other layers with regards to the specific architecture. The training time of the feed forward has a lower comparison to the recurrent type. However, the recurrent type has better memory for recalling previous events. In this experiment, the MLP classifier was used [11].

Radial Basis Function networks operate by implementing the regularization theory for problems that are ill-conditioned. For problems involving a map that changes input into desired classifications, the use of only a few samples may be problematic. Reducing the potential of error between the output of the network and the response that is desired is a way of addressing this problem. By processing mobile network cells in the hidden layer and then putting them over the input regions, the positions for each of the radial basis functions and the variances can be estimated using linear weights. The unsupervised learning rule of k-nearest neighbour is used in order to estimate the centers and widths. The size of the input space is obtained from the input data [4].

Estimating the Expected Costs of KPI Alarms Metrics

Classification Cost Matrix. The most complete way to provide the information about misclassification costs may be a Cost Matrix (also known as loss matrix), which indicates the costs for correct and incorrect classifications. Corresponding KPI alarms might affect the optimisation rules and the causes of network malfunctions based on previous research [12][13]. Given the cost matrix, an instance should be classified into a class that generates the minimum expected cost for four classes of KPI alarms correspondingly {a, b, c, d}. The expected cost of classifying an instance x into class can be expressed as a 4-class problem with the associated cost matrix given in Table 1 [3]:

Table 1. Network Cost Matrix

Cost Classification	Predicted Class	Actual Class			
		Traffic Costs	-2.5	4	4.5
Transmission Costs		2.1	-3.5	5	0
Infrastructure Costs		1.2	1.3	-4	1
Maintenance Costs		0.9	1.1	2.1	-3

The diagonal of the network cost matrix shows the costs for correct classification, all of which are negative values. This is normal because an appropriate classification has benefits rather than costs. The use of this type of network cost matrix to create classifiers that minimise the prediction cost as opposed to the prediction error has been used in some learning systems. However, it is also common for a learning system to be used that is not cost-sensitive. In this case, the class distribution of the training data set would be modified so that a classifier can be obtained that adjusts to the cost matrix and the class distribution [14].

Cost Algorithm Methodology: Measuring Performance of Individual KPI's

We have developed cost models that estimate the current cost of using weka cost sensitive learning classifier. The model takes into account the volume of transmitted calls and traffic of different services, as well as the infrastructure and maintenance operation costs that could be used to efficiently provide any particular service. Bayesian networks classifier optimised by classification accuracy may have higher misclassification cost for imbalanced classification problem. The cost-sensitive learning method is aimed to minimise classification costs [15].

An algorithm has been developed to estimate the network cost. Algorithm rules are developed for each of the following types of optimisation costs: analogue electronic switching; remote analogue switching; digital switching; remote digital switching; digital subscriber line carrier; copper loops and interoffice circuits; and t-carrier. Each of these costs are associated with specific services that are provided, such as a digital subscriber line or a business network connection. As each telecommunications company generally provides different types of services to different companies, it was important to attempt to define the costs associated with each of these services and the operations that are performed to provide them. At the same time, analogue technology was used for the stand-alone local network, while digital facilities were used for toll and enhanced services. Once again, the idea was to separate and define the costs based on the types of services provided to either home customers or business customers [2].

The optimisation algorithm has the ability to optimise the cost parameters based on the posterior probabilities that are determined by the neural network for the purpose of performance maximization. The class cost vector can be optimised based on the specific objective function in order to achieve the optimum classification performance. In this process, evaluation measures have an important role in terms of assessing the classification performance, as well as guiding the modeling of the classifier. With imbalanced datasets, the evaluation metric is able to take into account the imbalance that is present. The average accuracy is not an evaluation metric [1].

There are several costs involved in our cost optimisation prediction system. We use the key performance indicators alarms to provide cost-effective monitoring of network performance improvement. Our algorithm employs a tree structure to extract particular features for particular inputs, and we refer to it as the Cost-Sensitive Tree of Classifiers (CSTC). We begin by introducing foundational concepts regarding the CSTC tree and derive a global cost term that extends to trees of classifiers and then we relax the resulting loss function into a well-behaved optimisation problem [16].

From an algorithmic standpoint understanding pure cost optimisation is a major stepping stone for the more general criteria of cost-performance improvement that we plan to do in future work [16].

Cost Optimisation Algorithm:

Input data:

Training data sets D of cell data and KPIs, network faults based on Dropped call rate, Calls success rate and Traffic rate and Handover failures KPIs alarms $\{(DCR_{ji}, CSSR_{ji}, TR_{ji}, HOF_{ji})\}$ $n i=1,$

Initialise the cell data population encoding the randomly selected majority cost class examples $\{X_1, X_2, X_3, X_4\}$ for each individual set of data and KPIs.;

Traditional Rescaling of cost optimisation values:

Training set D with n classes of network costs, cost matrix
Network Matrix(cost)- Generate the matrix for optimisation costs

$COST_{tr}^i$	$COST_{inf}^i$
$COST_{trsmi}^i$	$COST_{mni}^i$

Output 1: Learning C45 algorithm trained the classifiers of training data and evaluate the network costs for each network cell individually.

Output 2: Builds a Cost Sensitive Naïve Bayes model on the KPI training data based on cost matrix

Cost classification Algorithm:

```

if DCR > d and CSSR <f while TR < e
  then transmission cost
  else if Rank CSSR <f or TR >e while HOF > c
    then traffic cost
  for KPI Alarm = “Warning class” while HOF> c
    then infrastructure cost
  for KPI Alarm = “Norm” class
    then maintenance cost
  
```

- Classify the training data based on the algorithm and evaluate costs considerations for each network cell.
- Generate cost classes output with the trained neural network.

Output 3: Builds a Neural Network using MPL, RBF and Multilayer Perceptron models on the training data for network costs..

With every class of cost decided its own fault diagnosis conditions, based on the fault diagnosis algorithm in our technical troubleshooting scheme that classifies the network faults in critical, warning or norm when the cell is optimised [12]. From the cost sensitive learning applications and the cost classification algorithm, four types of optimisation costs are generated. A low cell drop rate of between 4 and 2 was chosen, along with a low call success rate of between 90% and 98% in order to achieve an optimised cost related to transmission. The handover rate and traffic rate provide a clear and accurate measure of the conditions related to traffic costs on the network. The KPIs faults serve as warning faults and a high rate of handover failures need to be checked as infrastructure issues. It should also be noted that for the cells in which none of the network faults were valid, the algorithm would provide an estimate of the hypothetical average for the maintenance and operation costs.

The purpose of the algorithm is to generate an output that would show the results of the four different cost classes and evaluate the relations between the KPI alarms and the network costs considerations to assist financial predictions.

Performance Evaluation with Simulation Data

An evaluation was performed for the fault detection capability of the anomaly detection method that was previously explained in order to determine the suitability of the technique for telecommunication networks. This evaluation was conducted with network troubleshooting data. The simulation involved four classes of network faults, to which the anomaly detection method was applied [18] [12]. The algorithm consisted of a tree structure in order to extract specific inputs. The decision tree algorithm referred to as cost-sensitive classifiers. The J48 concepts were introduced into the algorithm, which allowed for the creation of global cost terms. The loss function that was created was relaxed into an optimisation problem [16].

Cost Evaluation Methodology and Accuracy

Four data sets of costs with computation and storage resources were used. Graphs with threshold curves from 0 to 1 and benefits from 1 to 2100 were created to examine the cost optimisation sensitivity for the four classifiers of the cost-benefit analysis as shown in Figure 3 [17].

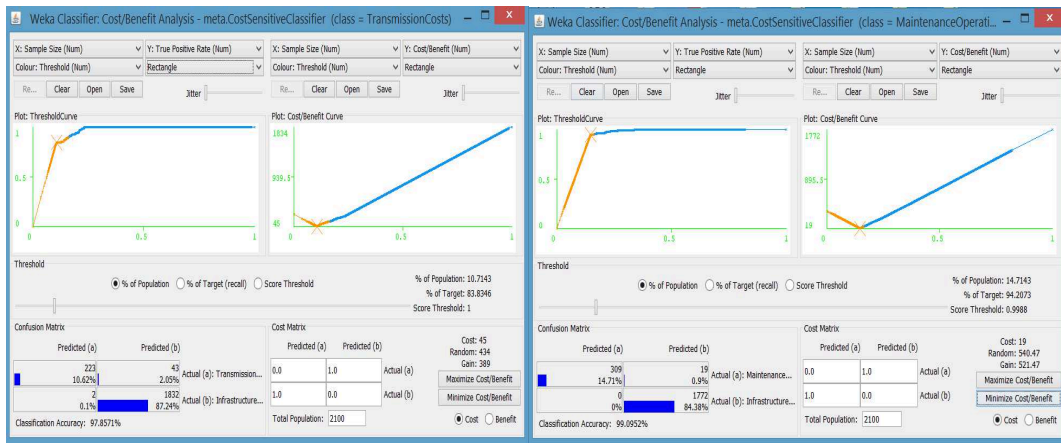


Fig. 1 Transmission and Maintenance cost benefit evaluation

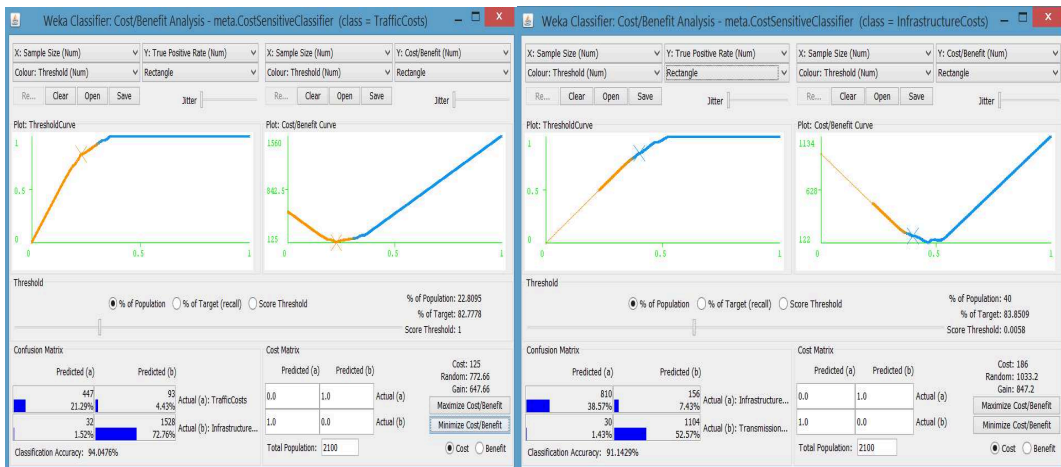


Fig. 2 Traffic and Infrastructure cost benefit evaluation

The result of the transmission cost benefit evaluation shown in Figure 1 has an accuracy rate of 97.85% and a minimum transmission cost of 45. In addition, the results show that problematic cells due to transmission are the 10.71% of an 83.83% target rate. The score set was 1 for threshold testing with the Naïve Bayes and J48 classifiers. The maintenance & operation cost evaluation results shown in Figure 1 have an accuracy rate of 99% with a minimum cost of 19. The

results also show that 14.71% of 2100 cells were problematic, along with a maintenance target of cost benefit rate of 94.2% and a score of 0.99 for threshold testing with Naïve Bayes and J48 classifiers.

The traffic cost evaluation results shown in Figure 2 provide an accuracy rate of 94% at a minimum cost of 125, which is too high for a financial plan. The results also show that 22.8% of the cells were problematic, and need to be reviewed and a traffic cost target rate of 82.7% with a 1 score set for threshold. The real traffic cost derived from the classifier is 93, with a predicted cost of 447.

The infrastructure cost benefit evaluation results demonstrate an accuracy rate of 91.41% at a minimum cost of 186. The results also show that 40% of the cells were problematic, with a target rate of 83.5% and a 0.0058 score set threshold. The actual traffic cost derived from the classifier is 158, with a predicted cost of 810.

Neural Networks Methodology and Accuracy

Artificial Neural Networks (ANN) are comprised of simple elements that operate in parallel to each other. The result is a multilayer perceptron (MLP) ANN that is fully connected with two hidden layers. The input layer was put into a high dimensional first hidden layer that allowed for effective features to be chosen. Then, for the output activation function, a nonlinear hyperbolic tangent function was used along with other activation functions in order to achieve the best results. The overall result was a nonlinear transformation in the network [11].

The architecture of Artificial Neural Networks (ANN) can be classified as feed forward or recurrent type. Each of four cost classes can be structured in different configurations in terms of the KPIs and the network locations. A feed forward network is a network in which the output of one layer is connected to the input of the next layer through a synaptic weight. The recurrent type can have at least once feedback connection between neurons with the same layer or from other layers based on the network architecture [11].

In this experiment, Multilayer Perceptron classifier was used that was based on the backpropagation learning algorithm to the classifier instances. The network was created by an MLP algorithm. The network could be also monitored and modified during the training. The nodes in the network were sigmoid with the exception of when the class was numeric. When the class was numeric, the output nodes became threshold linear units. The backpropagation neural network is a network comprised of simple processing elements that work together to produce some complex output. The backpropagation algorithm performs the learning on a multilayer feed-forward neural network. The algorithm learns a set of weights for a prediction of the class label of tuples. A multilayer feed-forward neural network is comprised of an input layer, at least one hidden layer, and an output layer [19].

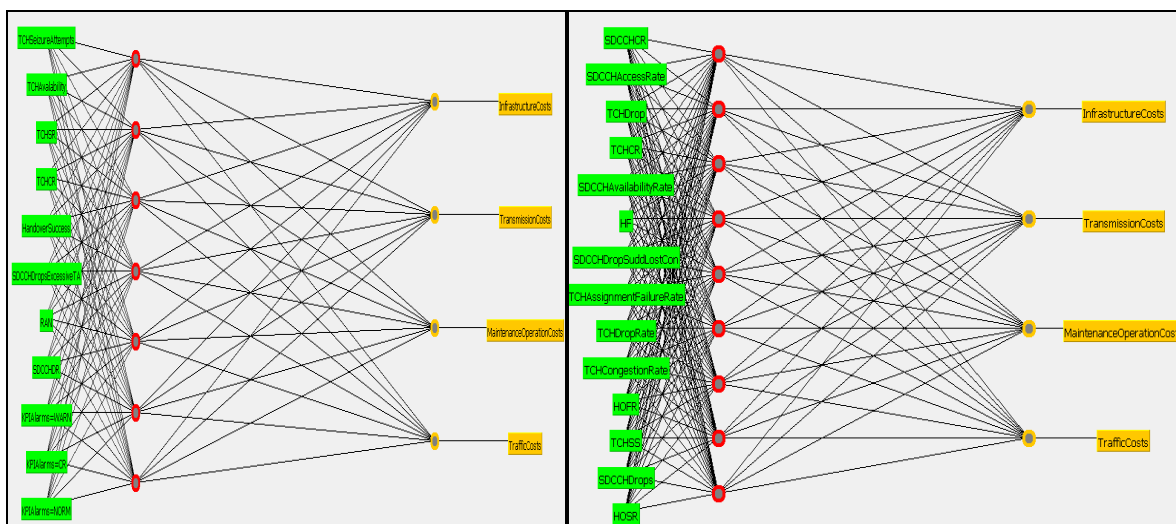


Fig.3 Neural network cost optimisation models for selected sets of KPIs

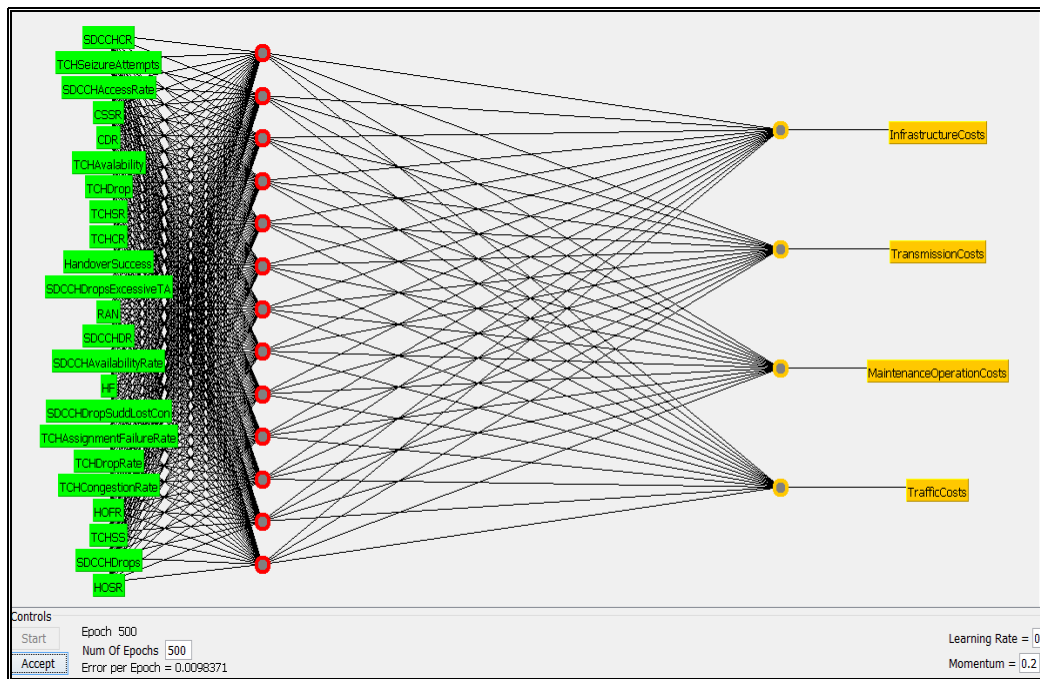


Fig.4 MLP parallel outputs for the final cost optimisation model

Figure 3 shows the multilayer feed-forward neural network for the network costs that was created. The neural network model used a feed-forward multilayer perceptron (MLP) classifier that defined the KPIs that were part of the different classification costs. In the first experiment that was performed, two different sets of KPIs were selected in order to identify the hidden values of the technical areas. In the second experiment, the Standalone Dedicated Control Channel Congestion Rate (SDCCHCR) variations, traffic channels dropped rates (TCHDR), and Handover success (HOSR) variations were selected. The goal was to connect the categories of the network costs with the input layers that were related to the KPIs for network performance. Based on the results that were achieved, the four NN cost models were able to adapt to the variations of the optimisation costs. The ANN model that was designed for financial predictions in the mobile networks [11].

A completely connected multilayer perceptron (MLP) ANN was used with 21 sets of KPIs shown in figure 4. The reason for this was because of the dynamic nonlinearity that was associated with the RF traffic pattern, as well as the random interfering signals from artificial and natural sources. The input layers were placed in the hidden layers in order to achieve appropriate selection of features. The action functions used in the 14 hidden layers were nonlinear hyperbolic tangent functions. In addition, a linear symmetric straight line was used for the output activation function [11].

Experimental Results

The results of the network cost model that was performed in this investigation allow for the conclusion that the simulation allowed for the effective evaluation of the cost of the network faults, as well as the ability to reduce the total misclassification cost. Previous research showed that it was possible to measure the distribution and the network faults as a means of identifying the cause of those faults [2]. In addition, the results of this investigation also allow for the conclusion that it is possible to minimise expected costs with the use of J48 rules as a means of minimising the expected operation costs.

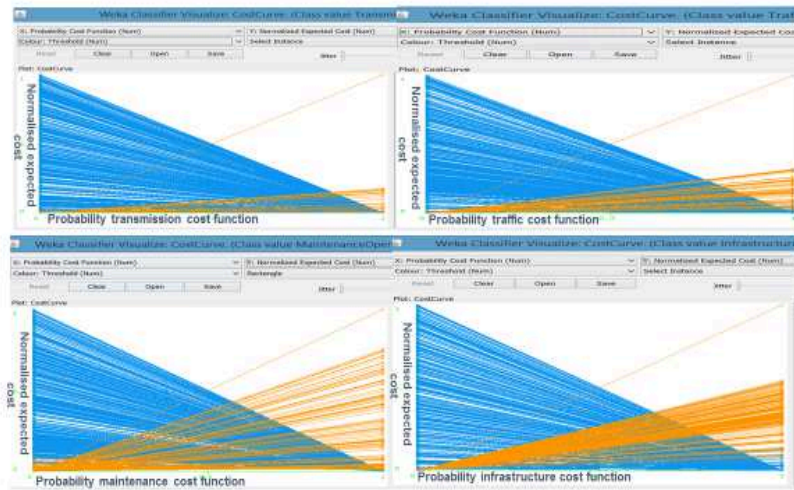


Fig.5 Probability cost in comparison to normalised expected costs

The output model on the posteriori probability of the class is shown in Figure 5 with the graphs of the four cost functions. The top two graphs show the transmission and traffic cost evaluation models, while the bottom two graphs show the maintenance and infrastructure cost evaluation models. The results shown such as TransmissionsClass (0, 0.1), TrafficClass (0, 0.3), MaintenanceClass (0.1, 0.7) InfrastructureClass (0, 0.5). The classes of Transmission costs and Traffic costs show the higher accuracy and precision without misclassified attributes at the final matrix. The cost curve indicates a higher sufficient level for the infrastructure and maintenance expected costs.

Table 2. Experimental Results on Cost-Sensitive Neural Networks Methods.

Results	BAYESNET	J48	RBF	MPL
Correctly Classified Instances	95.2%	100%	97.9%	99.9%
Kappa Statistics	0.93	1	0.96	0.99
Total Cost	-5609	-6278	-6124	-6274
Average Cost	-2.671	-2.989	-2.916	-2.988
Mean absolute error	0.027	0	0.063	0.0019
Root mean squared error	0.1445	0	0.11	0.013
Relative absolute error	7.91%	0	18.57%	0.55%
Root relative squared error	35%	0%	28.75%	3.17%

The accuracy rate of the Bayes network results shown in table 2 was 95.2%, with a total cost of -5609 and an average cost per cell of -2.674. The transmission and traffic costs show a higher rate of transmission and precision with regards to misclassified network cells. The multi-class classification of the KPIs also show the potential causes of the troubleshooting network issues, which is based on four cost parameters that are relevant to network optimisation.

The meta cost sensitive classifiers J48 results show an accuracy rate of 100%, with a total cost of -6278 and an average optimisation cost of -2.9895 per cell. Overall, what is shown is that the algorithm that was tested in this investigation reduces the misclassification cost.

The level of accuracy and generalization of ANN is primarily based on the initial weights and biases, learning rate, moment constant, training data and the network architecture. In this experiment, the learning rate and the momentum were maintained at 100% accuracy for J48, 99.95% for MPL, 97.9% for RBF, and 95.2% for Baynet. The initial weights and biases changed through the use of differential evolutionary algorithm variants [11].

The efficiency of the multilayer perceptron (MLP) classification model was found to be efficient in 99.9% of the instances after evaluating 17 Sigmoid Nodes, as well as after individually analysing the thresholds of all of the values of 25 KPIs. The MLP results showed a lower total cost of -349.9, and an average optimisation cost of -1.738 for each cell. The results of the RBF classifier model showed a reduced total cost of -6124, and an average optimisation cost of -2.9162 for each cell.

Figure 6 shows the output weight for the KPIs from the hidden units for the multilayer perceptron classifier model, as well as the variations in the output weights of the KPIs. Operators can update the output weights so that the hidden units can change cost classes based on network performance.

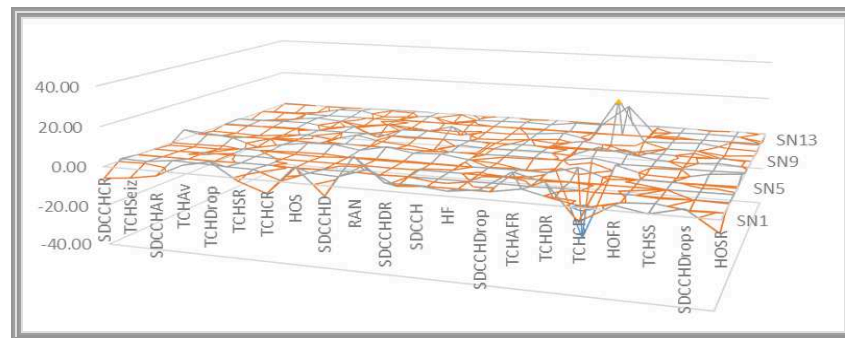


Fig.6 Multiple KPIs Output Weights

The hidden layer in the model acted as a feature detector so that during the training, the KPIs features were learned. In situations in which the abnormal values were shown in the hidden values, they were checked. For example, the Handover Success Rate (HOSR) and the Traffic Channel Congestion Rate values had to be checked in order to ensure accurate cost optimisation.

Conclusions

Having determined risk levels for all the projects in the study, we want to understand the impact of optimisation cost modelling in evaluating fault prediction models. More precisely, we want to know how well the models perform when misclassification cost is taken into account. Cost-sensitive learning solutions allow for the ability to reduce misclassification costs with the benefit of cost optimisation. What is demonstrated is that the binary foundation of many machine learning problems and algorithms need to be extended to have the ability to address multi-class classification problems [2].

The model put forward in this investigation provides the ability to determine the bandwidth cost that can be achieved through network troubleshooting management. In this way, network performance and cost are optimised at the same time. Meta cost classification techniques are effective and accepted tools for pattern recognition and categorisation. The cost sensitive classifiers are able to obtain a high level of accuracy while supplying an interpretable model for the end cost evaluation of optimisation faults.

Minimization of the cost function makes it possible for the neural network to both forecast and teach optimal cost decisions. The result is that the optimisation model process is more realistic because specific KPIs are selected based on the selected costs in question. The aspects of finding the optimal service levels for the network become part of the NN training process. This leads to predictions of cost minimum without the need for further calculations.

It is important to understand, however, that the network has to be trained several times in order to be able to find the combination of inputs that have the lowest error. In this experiment, genetic algorithms were used on each layer in RBF with different network architecture. Furthermore, the decision was made for online learning for updating the weight of the hybrid neuro-genetic RBF following the presentation of each exemplar [4].

On a broad level, determining an appropriate technique for valuation of fault detection and budgeting will improve the overall quality of services, especially for customers who are willing to

pay more for a higher quality of service. This technique can be used to address cost optimisation and finance decisions, define annual budgeting and planning efforts, provide support for management planning processes, and allow for the production of pricing and valuation information for customers. However, it is necessary to change and modify the criteria of cost-performance optimisation based on the different financial evaluations so that accurate predictions are achieved for a given network.

References

- [1] P. Cao, D. Zhao and O. Zaiane, Measure optimized cost-sensitive neural network ensemble for multiclass imbalance data learning. 2013 13th International Conference on Hybrid Intelligent Systems. (2013) 35-40.
- [2] C. Charnay, N. Lachiche and A. Braud, Pairwise optimization of bayesian classifiers for multi-class cost-sensitive learning, 2013 IEEE 25th International Conference on Tools with Artificial Intelligence. (2013) 499-505.
- [3] J. Kim, K. Choi, G. Kim and Y. Suh, Classification cost: An empirical comparison among traditional classifier, cost-sensitive classifier, and metacost, Expert Systems with Applications. 39, 4 (2012) 4013-4019.
- [4] N. Loukeris. Radial basis functions networks to hybrid neuro-genetic RBFNs in financial evaluation of corporations. 12 WSEAS International Conference on Computers. (2008) 812-819.
- [5] M. Adler, R. K. Sitaraman and H. Venkataramani, Algorithms for optimizing the bandwidth cost of content delivery, Computer Networks. 55, 8 (2011) 4007-4020.
- [6] D. Gabel, An application of stand-alone costs to the telecommunications industry, Telecommunications Policy. 15,1, (1991) 75-84.
- [7] V. Lopez, S del Rio, J. M. Benitez and f. Herrera, Cost-sensitive linguistic fuzzy role based classification systems under the MapReduce framework for imbalanced big data, Fuzzy Sets and Systems. 258 (2015) 5-38.
- [8] P. N. Um, L. Gille, L. Simon and C. Rudelle, A Model for calculating interconnection costs in Telecommunications. Washington, D.C.: The World Bank. (2004).
- [9] B. Haider, M. Zafrullah and M. K. Islam, Radio frequency optimization & QoS evaluation in operational gsm network, Proceedings of the World Congress on Engineering and Computer Science. (2009) 393-398.
- [10] C. K. Walgampaya and M. Kantardzic, Cost-sensitive analysis in multiple time series prediction, Conference on Data Mining. (2006) 17-23.
- [11] S. Iliya, E. Goodyer, J. Gow, J. Shell and M. Gongora. Application of artificial neural network and support vector regression in cognitive radio networks for RF power prediction using compact differential evolution algorithm, Proceedings of the Federated Conference on Computer Science and Information Systems. 5 (2015) 55-66.
- [12] E. Rozaki, Design and implementation for automated network troubleshooting using data mining, International Journal of Data Mining & Knowledge Management Process (IJDKP). 5 (2015) 9-27.
- [13] E. Rozaki, Clustering optimisation techniques in mobile networks, International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC). 4, 2, (2016) 22 – 29.
- [14] C. Ferri-Ramirez, P. Flach and J. Hernandez-Orallo, Multi-dimensional roc analysis with decision trees, Technical Report. (2002) 1-36.

- [15] Y. Gao and J. Wang, Active learning method of bayesian networks classifier based on cost-sensitive sampling, 2011 IEEE International Conference on Computer Science and Automation Engineering. (2011) 233-236.
- [16] Z. Xu, M. J. Kusner, K. Q. Weinberger, M. Chen and O. Chapelle, Classifier cascades and trees for minimizing feature evaluation cost, Journal of Machine Learning Research. 15 (2014) 2113-2144.
- [17] L. Zeng, B. Veeravalli and X Li, SABA: A security-aware and budget-aware workflow scheduling strategy in clouds, Journal of Parallel and Distributed Computing. 75 (2015) 141-151.
- [18] T. Kihara, N. Tateishi and S. Seto, Evaluation of network fault-detection method based on anomaly detection with matrix eigenvector, 2011 13th Asia-Pacific Network Operations and Management Symposium. (2011) 1-7.
- [19] R. Arora and Suman, Comparative analysis of classification algorithms on different datasets using WEKA. International Journal of Computer Applications. 54, 13 (2012) 21-25.

Financial Predictions Using Cost Sensitive Neural Networks for Multi-Class Learning

10.4028/www.scientific.net/AEF.16.104

DOI References

- [1] P. Cao, D. Zhao and O. Zaiane, Measure optimized cost-sensitive neural network ensemble for multiclass imbalance data learning. 2013 13th International Conference on Hybrid Intelligent Systems. (2013) 35-40.
10.1109/his.2013.6920500
- [2] C. Charnay, N. Lachiche and A. Braud, Pairwise optimization of bayesian classifiers for multiclass cost-sensitive learning, 2013 IEEE 25th International Conference on Tools with Artificial Intelligence. (2013) 499-505.
10.1109/ictai.2013.80
- [3] J. Kim, K. Choi, G. Kim and Y. Suh, Classification cost: An empirical comparison among traditional classifier, cost-sensitive classifier, and metacost, Expert Systems with Applications. 39, 4 (2012) 4013-4019.
10.1016/j.eswa.2011.09.071
- [5] M. Adler, R. K. Sitaraman and H. Venkataramani, Algorithms for optimizing the bandwidth cost of content delivery, Computer Networks. 55, 8 (2011) 4007-4020.
10.1016/j.comnet.2011.07.015
- [6] D. Gabel, An application of stand-alone costs to the telecommunications industry, Telecommunications Policy. 15, 1, (1991) 75-84.
10.1016/0308-5961(91)90045-d
- [7] V. Lopez, S del Rio, J. M. Benitez and f. Herrera, Cost-sensitive linguistic fuzzy role based classification systems under the MapReduce framework for imbalanced big data, Fuzzy Sets and Systems. 258 (2015) 5-38.
10.1016/j.fss.2014.01.015
- [8] P. N. Um, L. Gille, L. Simon and C. Rudelle, A Model for calculating interconnection costs in Telecommunications. Washington, D.C.: The World Bank. (2004).
10.1596/0-8213-5671-2
- [10] C. K. Walgampaya and M. Kantardzic, Cost-sensitive analysis in multiple time series prediction, Conference on Data Mining. (2006) 17-23.
10.1109/ijcnn.2006.247298
- [12] E. Rozaki, Design and implementation for automated network troubleshooting using data mining, International Journal of Data Mining & Knowledge Management Process (IJDKP). 5 (2015) 9-27.
10.5121/ijdkp.2015.5302
- [15] Y. Gao and J. Wang, Active learning method of bayesian networks classifier based on costsensitive sampling, 2011 IEEE International Conference on Computer Science and Automation Engineering. (2011) 233-236.
10.1109/csae.2011.5952671
- [17] L. Zeng, B. Veeravalli and X Li, SABA: A security-aware and budget-aware workflow scheduling strategy in clouds, Journal of Parallel and Distributed Computing. 75 (2015) 141-151.
10.1016/j.jpdc.2014.09.002
- [18] T. Kihara, N. Tateishi and S. Seto, Evaluation of network fault-detection method based on anomaly detection with matrix eigenvector, 2011 13th Asia-Pacific Network Operations and Management Symposium. (2011) 1-7.

10.1109/apnoms.2011.6077024

[19] R. Arora and Suman, Comparative analysis of classification algorithms on different datasets using WEKA. *International Journal of Computer Applications*. 54, 13 (2012) 21-25.

10.5120/8626-2492