EAI Endorsed Transactions
on Intelligent Systems and Machine Learning Applications

EAI.EU
Research Article

Optimising performance indicators in the telecommunications sector

Fiston Chrisnovic Balanganayi Kabutakapua [1*], Pierre Kafunda Katalay[2], Simon Ntumba Badibanga[3], Eugène Mbuyi Mukendi[4]

[1, 2, 3, 4]Department of Mathematics, Statistics and Computer Science, Faculty of Science and Technology, University of Kinshasa, Kinshasa, D.R.Congo

Abstract

INTRODUCTION: This study analyses and predicts performance in the public telecommunication sector using neural networks on key performance indicators in the telecommunication sector.
OBJECTIVES: Although there are several key performance indicators in the telecommunications sector, we have selected a few and assessed their correlation with the variable to be predicted. In this study, we used supervised learning based on a multi-layer neural network.
METHODS: The algorithm used in this study is the retro propagation algorithm because of its simplicity and accuracy of estimation.
RESULTS: The results show that the selected indicators, including accessibility, maintainability, satisfaction, network operating cost and availability, explain more than 80% of the performance in the telecommunications sector, and the area of the ROC curve is equal to 0.97, which means that the classifier is almost perfect. This is also justified by the sensitivity and specificity, which are close to 1 when observing the ROC curve and the confusion matrix. The classification error found from the confusion matrix is equal to 1%, which means that our model has very high accuracy.
CONCLUSION: The other indicators presented were not selected in the model because of their low correlation with the variable of interest and the difficulty of collecting the data.

Keywords: performance in the telecommunication sector, neural network on performance indicators, supervised learning, multilayer neural network, public telecommunication sector

Received on 15 February 2025, accepted on 30 May 2025, published on 06 June 2025

doi: 10.4108/eetismla.8717

1. Introduction

Digital transformation has created a number of changes in industries and highlighted the importance of data and its analysis. This is particularly true for the telecoms sector, which relies heavily on different types of data analysis to define key performance indicators that impact everything from operations and network performance to customer experience and overall profitability. Data analysis is the process of analysing large sets of data to identify trends and derive valuable insights. Key Performance Indicators (KPIs) determine how companies measure their success against defined objectives. Today's telecom service providers need a way to constantly monitor themselves if they hope to reduce expenses and streamline revenue streams, and that's where KPIs can help.

The definition of data analysis is open to interpretation depending on the industry you are looking at. For the telecoms sector in particular, the meaning of data analysis includes uncovering the truth about a telecoms company's performance in the marketplace. This data provides insight

*Corresponding author. Email: fistonbalang@gmail.com

into all aspects of the business and how it operates. This enables a telecoms company to identify any problems, issues or concerns, make continuous improvements and create opportunities for growth, impacting revenue and profitability.

Key Performance Indicators (KPIs) are used to measure a company's performance over a given period. How does data analysis fit in with these indicators? The answer is simple: data analysis provides the data needed to respond to the KPIs.

Data analysis uses software, machine learning and artificial intelligence to collect, store, clean, analyse and interpret raw data from all the different data streams linked to the operator's network. Key Performance Indicators to model the data in a way that makes sense for your business

Data on usage patterns for telecoms services can be analysed to forecast future demand, which can help operators plan capacity and allocate resources efficiently.

Data mining can be used to analyse network performance data to identify bottlenecks and quality of service problems, enabling operators to take corrective action.

The purpose of this guide is to identify, define and analyse the main indicators for analysing the public telecommunications sector. By "public telecommunications sector" we mean the telecommunications infrastructure and services offered to the general public. We will use machine learning based on neural networks to predict the level of performance on the basis of the key indicators we have selected.

1.1. Theory on telecommunication performance indicators

Key performance indicators provide a telecoms company with the information it needs to ensure that the network is working as it should and that customers are satisfied with their experience and services. These elements are ultimately the backbone of a telecoms company's business and guarantee its success, hence the importance of these indicators and the data analysis process.

Key performance indicators in the telecommunications sector include:

- Accessibility - This involves determining whether users can access a requested service and the quality of the service available when needed. This also includes technical elements relating to the success rate of the installation of the CRR, the success rate of the installation of the ERAB and the success rate of the installation of calls.
- Maintainability - Measures whether the network can hold up and deliver the promised network service to users. In this case, call drop rates and service call drop rates are monitored.
- Mobility - Measures network performance while users are on the move. This makes it possible to monitor the technical elements relating to the intra-frequency handover success rate, the inter-frequency handover success rate and the inter-RAT handover success rate (LTE to WCDMA).
- Integrity - Measures network quality, throughput and latency. This includes E-UTRAN IP throughput, DL IP throughput and E-UTRAN IP latency.
- cess a requested service and the quality of the service available when needed. This also includes technical elements relating to the success rate of the installation of the CRR, the success rate of the installation of the ERAB and the success rate of the installation of calls.
- Maintainability - Measures whether the network can hold up and deliver the promised network service to users. In this case, call drop rates and service call drop rates are monitored.
- Mobility - Measures network performance while users are on the move. This makes it possible to monitor the technical elements relating to the intra-frequency handover success rate, the inter-frequency handover success rate and the inter-RAT handover success rate (LTE to WCDMA).
- Integrity - Measures network quality, throughput and latency. This includes E-UTRAN IP throughput, DL IP throughput and E-UTRAN IP latency.
- Availability - Indicates whether the network is suitable or ready for use by users.
- Utilisation - Monitoring network capacity and utilisation.
- Average return per user - Indicates how much money the MNO site earns for each person using its service.
- Subscriber Acquisition Cost - Shows how much money is spent on acquiring new subscribers. Costs such as marketing, advertising, sales commissions and the cost of getting customers onto your network are tracked here.
- Tracking churn - This is linked to sales. If your business is losing customers, you need to look at new customer acquisition, your service and your costs (i.e. look at the root cause of why customers are leaving and solve it).
- Network operating cost - This is the cost of operating your network.
- Net Promoter Score - The NPS (Net Promoter Score) is a performance measure used to assess the likelihood of a customer recommending your company to a friend or colleague.

In addition to these performance indicators, we can add the following indicators for the customer section:

- Network availability ;
- Call quality ;
- The quality of the Internet connection ;
- Latency ;
- Stability of the Internet connection ;
  Service availability;
- Customer satisfaction.

Types of data on key performance indicators

a. Periodic data

Indicators should be collected at least once a year. Historical data should be kept to measure trends and forecast future demand. The use of the same accounting date (e.g. the end of the calendar year) is desirable to allow comparisons. Although this is not always possible due to differing national practices, at least all data should relate to the same period in order to increase the accuracy of calculations. For example, operating data and financial data should apply to the same period, to enable revenue per subscriber line to be accounted for accurately.

a. Demographic and macro-economic data

The list of indicators does not include demographic or macro-economic statistics. These data are needed to assess the penetration of the telecommunications network and the impact of telecommunications on the economy as a whole. Demographic and macro-economic statistics are regularly collected and disseminated by the national statistical services and by the ministries of the economy and finance, as well as by regional and international organisations (OECD, UN, World Bank, International Monetary Fund, etc.).

1.2. Machine learning techniques for analysing telecommunication data

Dataming techniques are widely used in the public telecommunications sector for the following reasons:

- Social network analysis: telecoms operators can use data mining to analyse communication patterns between subscribers, which can help them better understand network usage trends and improve services.
- Demand analysis: data on usage patterns for telecoms services can be analysed to forecast future demand, which can help operators plan capacity and allocate resources efficiently.
- Fraud detection: data mining can be used to detect patterns of fraud in communications, such as misuse of services or SIM card fraud.
- Customer segmentation: using clustering and segmentation analysis techniques, operators can group subscribers into homogeneous segments according to their needs, enabling them to personalise their offers and services.
- Churn prediction: by analysing customer behavioural data, operators can predict which subscribers are likely to churn and implement appropriate retention strategies.
- Product recommendation: by analysing customer usage patterns and preferences, operators can recommend personalised services or packages to improve the customer experience and increase revenues.

Data mining techniques for prediction include:
Some supervised learning techniques

The SVM
A Support Vector Machine is a supervised machine learning algorithm that can be used for classification and regression. SVMs are most commonly used in classification situations. SVMs are based on the idea of finding a hyperplane that best divides a dataset into two classes, as shown in the image below (Theodoros & Massimiliano, 2001).

0. Neural networks
1. A neural network is a mathematical model inspired by the functioning of the human brain, used in artificial intelligence to perform machine learning tasks. It is composed of several computational units (neurons) interconnected by weighted connections, and is capable of learning from data by adjusting these connection weights (Charu 2018).
2. A neural network is a system whose design was originally inspired schematically by the functioning of biological neurons, and which has subsequently moved closer to statistical methods. Neural networks are generally optimised using probabilistic learning methods. They are placed in the family of statistical applications, which they enhance with a set of paradigms that enable rapid classification (Kohonen networks in particular), and in the family of artificial intelligence methods, for which they provide a perceptual mechanism that is independent of the ideas of the person implementing it, and provides input information for logical reasoning.

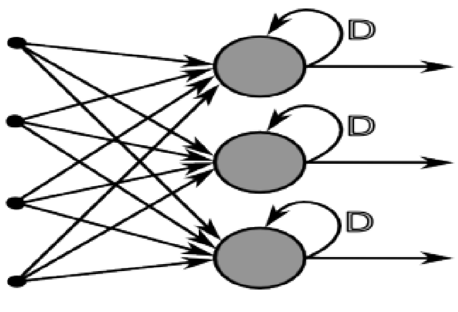

Figure 1: diagram of a neuron

The decision tree

Decision trees are classification rules that base their decision on a series of tests associated with the attributes, the tests being organised in a tree structure.
The internal nodes are called decision nodes. Each decision node is labelled with a test that can be applied to any description of an individual in the population. In general, each test examines the value of a single attribute in the description space. The possible responses to the test correspond to the labels of the arcs originating from this node.
Leaves are labelled by a class called the default class. Each internal node or leaf is identified by its position: the list of the numbers of the arcs that enable it to be accessed from the root. Decision trees are one of the most widely used non-parametric supervised learning methods in classification and regression. This is partly because of their algorithmic simplicity and partly because they are easy to interpret and explain the results generated. Decision trees are built using an algorithmic approach and can be visualised as a tree with rules that identify the ways in which a set of data can be split. The aim is to create a model that predicts the value of a target variable by learning the decision rules.

Example of a decision tree

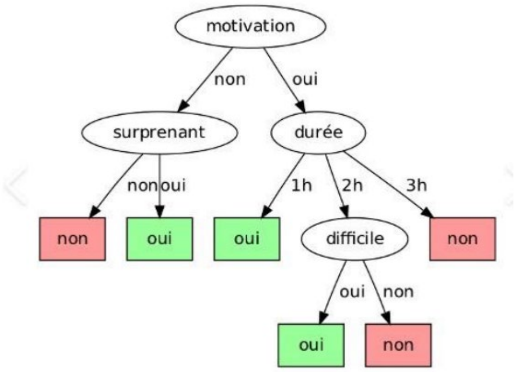

Figure 2: Decision tree

A decision tree is a decision-making tool that represents a set of choices in the graphic form of a tree. The various possible decisions are located at the ends of the branches (the "leaves" of the tree). They are reached on the basis of decisions made at each stage. The decision tree is a tool used in a variety of fields, including security, data mining and medicine. It has the advantage of being easy to read and quick to execute. It is also a representation that can be calculated automatically by supervised learning algorithms.

Decision trees are powerful and popular tools for classification and prediction (Hastie, op cit).

The K nearest neighbour
Nearest Neighbours is a machine learning technique and algorithm that can be used for regression and classification tasks. Nearest Neighbors examines the labels of a selected number of data points surrounding a target data point, in order to make a prediction about the class to which the data point belongs. K-Nearest Neighbors (KNN) is a conceptually simple but very powerful algorithm, and for these reasons it is one of the most popular machine learning algorithms ((Hastie, op cit). Let's take a deep dive into the KNN algorithm and see exactly how it works. Having a good understanding of how KNN works will enable you to appreciate the best and worst use cases for KNN.

A KNN algorithm goes through three main phases during its execution:

1. Set K to the chosen number of neighbours.
2. Calculates the distance between a supplied/test example and the dataset examples.
3. Sorting calculated distances.
4. Obtain the labels for the upper K inputs.

The K nearest neighbour
Nearest Neighbours is a machine learning technique and algorithm that can be used for regression and classification tasks. Nearest Neighbors examines the labels of a selected number of data points surrounding a target data point, in order to make a prediction about the class to which the data point belongs. K-Nearest Neighbors (KNN) is a conceptually simple but very powerful algorithm, and for these reasons it is one of the most popular machine learning algorithms ((Hastie, op cit). Let's take a deep dive into the KNN algorithm and see exactly how it works. Having a good understanding of how KNN works will enable you to appreciate the best and worst use cases for KNN.

A KNN algorithm goes through three main phases during its execution:

1) Set K to the chosen number of neighbours.
2) Calculates the distance between a supplied/test example and the dataset examples.
3) Sorting calculated distances.
4) Obtain the labels for the upper K inputs.
5) Returns a prediction about the test example.

In the first step, K is chosen by the user and tells the algorithm how many neighbours (how many surrounding data points) to take into account when making a judgement about the group to which the target example belongs. In the second step, note that the model checks the distance between the target example and each example in the dataset. The distances are then added to a list and sorted. Next, the sorted list is checked and the labels of the top K elements are returned. In other words, if K is set to 5, the model checks the labels of the first 5 data points closest to the target data point. When rendering a prediction on the target data point, it matters whether the task is a regression or classification task. For a regression task, the average of the top K labels is used, whereas the top K labels mode is used in the case of classification.

The exact mathematical operations used to perform KNN differ depending on the distance metric chosen. If you would like to know more about how the metrics are calculated, you can find out more about some of the most common distance metrics, such as Euclidean, Manhattanand Minkowski.

Bagging

Proposed by Breiman in 1996, the principle of Bagging is to combine the results of several models (for example, several decision trees) to obtain a generalised result.

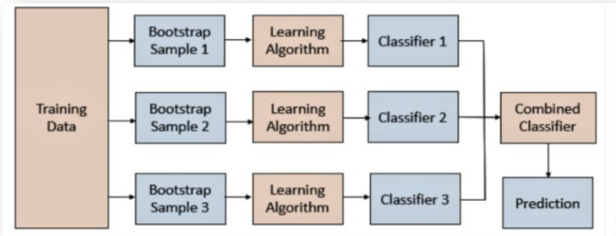

Figure 3: Principle of Bagging (Bootstrap Aggregation) [3].

- Bagging-based algorithms :
  - Bagging meta-estimator
  - Random forest
- Advantages and disadvantages of bagging :

  Random forest is one of the most popular bagging algorithms. Bagging has the advantage of allowing many weak learners to combine their efforts to outperform a single learner. It also helps to reduce variance, eliminating model over-fitting in the procedure.
  - One disadvantage of bagging is that it introduces a loss of interpretability into a model.
  - The resulting model can be subject to numerous biases when the appropriate procedure is ignored.
  - Although bagging is very accurate, it can be computationally expensive and this may discourage its use in some cases.

Stacking:

Stacking is an ensemble learning technique that uses the predictions of several models (e.g. a decision tree) to build a new model. This model is used to make predictions on the whole data set. It uses a meta-learning algorithm to learn how best to combine the predictions of two or more basic machine learning algorithms. The advantage of Stacking is that it can exploit the capabilities of a range of models performing well on a classification or regression task and make predictions that perform better than any single model in the set.

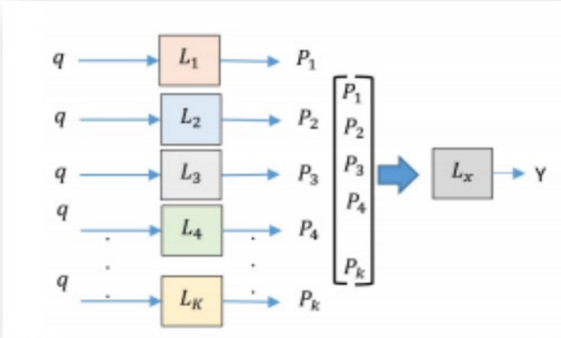

Figure 4: Classifier based on general stacking [6].

The aim of this chapter was to give a general overview of machine learning by detailing the principles and methods of supervised and unsupervised learning. We will move on to the second chapter where we will discuss the K-nearest neighbour method in more detail, presenting its advantages and limitations, and then highlight the parameters for the performance of our model.

2. Methodological aspects

a. Data source

Our target population is all the customers and staff of the mobile network reached, aged between 18 and 64. Each sales outlet (in this study we selected three) is a stratum representing a homogeneous group of statistical units.

b. Drawing the sample

The size of the sub-sample$n_i = \frac{N_i}{N} * n$ with N_i the size of stratum i, i.e. the number of customers and staff in outlet i, N the size of the population, i.e. all the customers and staff and n the size of the sample.

N_1 The size of the stratum of customers and staff in the first outlet,N_2 The size of the stratum of customers and staff in the second outlet andN_3 The size of the stratum of customers and staff in the third outlet.n_i is the sample size in stratum i. The sample size is found using the formula: $n = \frac{4Z^2_{\frac{\alpha}{2}}p(1-p)}{\varepsilon^2}$

P is the proportion of mobile network users we have chosen in the DRC that is unknown in 2024. As p is unknown, the literature requires us to take 0.5. ε is equal to 0.05 which is the error term and$Z_{\alpha/2}$ is the critical value read from the normal distribution table. THE SAMPLE SIZE IS$n = \frac{4Z^2_{\frac{\alpha}{2}}p(1-p)}{\varepsilon^2} = 1536$. So to build this sample, we used the proportionality rule to find the sample size in each outlet by the formula $n_i = \frac{N_i}{N} * n$then $\sum_{i=1}^{3} n_i = n$

c. Presentation of the algorithm used

To implement our neural network with R software, we will use the back-propagation algorithm. The choice of this algorithm is justified by the following advantages:

- Back propagation is fast, simple and easy to programme
- It has no parameters to set apart from the number of inputs.
- This is a flexible method because it requires no prior knowledge of the network.
- This is a standard method that generally works well
- It is not necessary to specifically mention the characteristics of the function to be learned.

The retro propagation algorithm is as follows:

1. Inputs X, arrive via the pre-connected path
2. The input is modelled using real weights W. The weights are generally selected at random.
3. Calculate the output for each neuron from the input layer, through the hidden layers, to the output layer.
4. Calculate the error in the outputs
5. Return from the output layer to the hidden layer to adjust the weights to reduce the error.

The output function is given by the expression: $output = 0\ si\ \sum w_i X_i \leq t_o et\ output = 1\ si\ \sum w_i X_i > t_o$. The output of a neuron with inputsX_i and weightsw_i and bias b is given by the function :

$output = \frac{1}{1+\exp\left(-\sum(w_i e_i)-b\right)}$ and it is a sigmoid neuron.

3. Presentation of Results

a. Data scaling

This step is very important when the data do not have the same scale, as it can lead to meaningless results. Several methods are used to scale data, such as: min-max normalisation, Z-score normalisation, median and tan-h estimators. In this study we used min-max normalisation.

b. Data sampling

We now divide the data into a training set and a test set. The training set shows the relationship between the dependent variable and the independent variables, while the test set shows the performance of the model. The index variable is used when fitting the neural network to create training and test data sets.

c. The neural network model

Let's now build a neural network on our data. The language we used to build our neural network is R, thanks to its neuralnet () library.

The algorithm we used is the one mentioned earlier, the retro propagation algorithm. For the first output, the error is 0.00925 and for the second output, the error is 65.8107, so the first output is the best because the error is small.

The architecture is as follows:

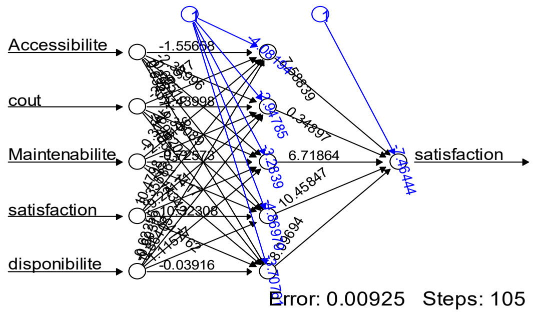

Figure 5: The neural network with weight estimation

1. Quality of the estimate

First, we will assess the quality of the estimate using the coefficient of multiple determination. The coefficient of multiple determination is nothing more than the square of the correlation coefficient between the observed variable and the predicted variable. In our case$R^2 = 0{,}89$ which justifies the good quality of the estimate since the value is greater than 0.5.

We then evaluate the quality of the estimate using the ROC curve. The surface is interpreted as follows:

- 1.0: Perfect classifier. The model classifies all instances perfectly;
- 0.9-1.0: Excellent classifier ;
- 0.8-0.9: Very good classifier ;
- 0.7-0.8: Good classifier ;
- 0.6-0.7: Fairly acceptable classifier ;
- 0.5-0.6: Poor performance ;
- 0.5: Random classifier.

The curve is as follows:

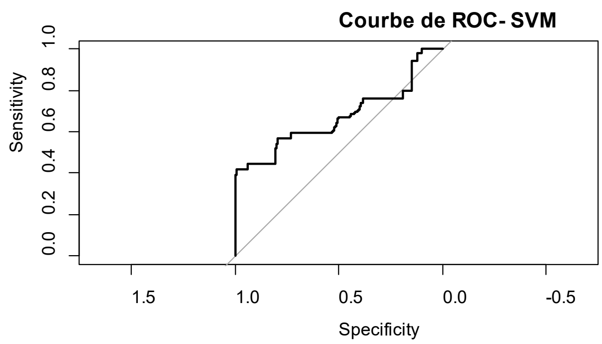

Figure 6: Rock curve

The area of the curve calculated using R software is equal to 0.97, which means that the classifier is almost perfect.

The other method for seeing the quality of the prediction is the confusion matrix. This allows us to calculate the estimation error and see the number of true positives, the number of false positives, the number of false negatives and the number of true negatives. In our case, it is the estimation error that interests us.

I.
Table 1: Classification (Confusion matrix)

Observed	Forecasts		
	Negative	Positive	Percentage correct
Negative	543	10	98,19%
Positive	4	542	99,26%
Overall percentage	99,26%	98,18%	98,73%
Growth method: Neural network Dependent variable : Satisfaction			

Source: Network quality survey

The model generates 542 true negatives, 543 true positives, 4 false negatives and 10 false positives. Let's calculate the misclassification error. The formula is$1 - tx\ bonne\ classification\ de\ classification$.

The correct classification rate is$\frac{sum(diag(M_{ii}))}{sum(M_{ij})} = \frac{1085}{1099} = 0{,}99$ withM_{ij} the elements of the confusion matrix (M). The misclassification error is therefore$1 - 0{,}99 = 0{,}01$ and1% , which means that our model provides high accuracy.

We can further increase the accuracy and efficiency of our model by increasing or decreasing the number of nodes and the bias in the hidden layers.

The strength of machine learning algorithms lies in their ability to learn and improve each time they predict an output. In the context of neural networks, this means that the weights and biases that define the connection between neurons become precise. This is why the weights and biases are selected in such a way that the output of the network is close to the true value for all the training inputs.

4. Discussion

The selected indicators clearly explain more than 80% of the performance in the public telecommunications sector. The hypotheses made on the influence of network accessibility, maintainability, availability, satisfaction and network operating costs on performance in the public telecoms sector are validated. However, the non-availability of data on other key performance indicators remains a limitation of this study. The collection of data on almost all the indicators would be better for improving the learning and explanation models. In all cases, a comparative analysis with other supervised learning techniques is better for improving estimation.

5. Conclusion

The objective of this study was to predict performance in the public telecommunications sector using key performance indicators. First we introduce the following. Next, we discuss the theory of key performance indicators to explain their relevance. Thirdly, we talk about data mining, presenting its techniques and its importance in the public telecommunications sector. Then we presented the method used for prediction and the data source. Finally, we presented the results of our study. The results showed that the selected performance indicators explain more than 80% of the performance in the public telecommunication sector. However, we suggested collecting data on the other indicators to further improve the model's performance.

References

[1] Alpaydin, E. (2014). Introduction to machine learning. MIT press.

[2] Alpaydin, E. (2020). Introduction to machine learning. MIT Press.

[3] "Applied Logistic Regression" by Hosmer Jr, Lemeshow, and Sturdivant, published in 2013.

[4] Babbie, E. (2016). The practice of social research. Cengage Learning.

[5] Cortes, C., Pregibon, D. Signature-based methods for data streams. Data Mining and Knowledge Discovery 2001; 5(3):167-182.

[6] Cortes, C., Pregibon, D. Giga-mining. Proceedings of the Fourth International Conference on Knowledge Department of Telecommunications. 1992. Indian Telecommunication Statistics 1991.

[7] Discovery and Data Mining; 174-178, 1998 August 27-31; New York, NY: AAAI Press, 1998.

[8] EUROSTAT. 1994. Telecommunication Indicators. Documento distribuido en el ICOBS/AS Working.

[9] Ezawa, K., Norton, S. Knowledge discovery in telecommunication services data using Bayesian network models. Proceedings of the First International Conference on Knowledge Discovery and Data Mining; 1995 August 20-21. Montreal Canada. AAAI Press: Menlo Park, CA, 1995.

[10] Fawcett, T., Provost, F. Adaptive fraud detection. Data Mining and Knowledge Discovery 1997; 1(3):291-316.

[11] Fawcett, T, Provost, F. Activity monitoring: Noticing interesting changes in behavior. Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 53-62. San Diego. ACM Press: New York, NY, 1999.

[12] Group Meeting, 12-13 de enero de 1994, Luxembourg. Federal Communications Commission. 1992.

[13] Han, J., Altman, R. B., Kumar, V., Mannila, H., Pregibon, D. Emerging scientific applications in data mining. Communications of the ACM 2002; 45(8): 54-58.

[14] Kaplan, H., Strauss, M., Szegedy, M. Just the fax-differentiating voice and fax phone lines using call billing data. Proceedings of the Tenth Annual ACM-SIAM Symposium on Data Mining In Telecommunications.

[15] Discrete Algorithms. 935-936. Baltimore, Maryland. Society for Industrial and Applied Mathematics: Philadelphia, PA, 1999.

[16] Klemettinen, M., Mannila, H., Toivonen, H. Rule discovery in telecommunication alarm data. Journal of Network and Systems Management 1999; 7(4):395-423.
[17] Mani, D. R., Drew, J., Betz, A., Datta, P. Statistics and data mining techniques for lifetime value modeling. Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 94-103. San Diego. ACM Press: New York, NY, 1999.
[18] Roset, S., Murad, U., Neumann, E., Idan, Y., Pinkas, G. Discovery of fraud rules for telecommunications-challenges and solutions. Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 409-413, San Diego CA. New York: ACM Press, 1999.
[19] Sasisekharan, R., Seshadri, V., Weiss, S. Data mining and forecasting in large-scale telecommunication networks. IEEE Expert 1996; 11(1):37-43.
[20] Weiss, G. M., Hirsh, H. Learning to predict rare events in event sequences. Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining. 359-363. AAAI Press, 1998.
[21] Weiss, G. M., Provost, F. Learning when training data are costly: The effect of class distribution on tree induction. Journal of Artificial Intelligence Research 2003; 19:315- 354.