# Coupling OLAP and Data Mining for Prediction

**Najlae KORIKACHE and Amina YAHIA**

BP 162 Oujda Principale
Faculté des sciences, Université Med I Oujda, Morocco

## Abstract

The cooperation of transactional databases and decision making systems resulted in warehouse concepts, multidimensional data bases (cube) and OLAP system. These concepts aim to implement effective solutions for the management and interactive analysis of large volumes of data and provide decision-making process. Integrating the prediction in OLAP environment is a topic that is experiencing a growing interest which opens several research paths.

This contribution addresses the problem of predicting from a data cube and consists of partitioning the original cube into dense sub-cubes as well as building and validating for each dense sub-cube a prediction model. It also consists of choosing (by the user) the cell to predict rather than the context of analysis and determining the sub-cube which contains the designated cell by the user and then predict the value of the cell through the prediction model of the sub-cube.

**Keywords**: OLAP, data mining, cube, sub data, sub dense cubes, prediction,

## 1 Introduction

In the competitive economic environment nowadays, information plays a crucial role in the daily business. The acquisition, analysis and use of information have become key strategic choices. The control of information is a crucial skill for any company wanting to establish itself in the forefront of its field. Given these requirements, large volumes of production data relating to the business of the company, have become real mine of knowledge. Great efforts were made to control large amounts of data and to extract potential knowledge from these data.

Data warehouses provided adequate and effective solution to the problem of storage and data management. A warehouse is a centralized database of large volumes of data, illustrated, organized by topic and consolidated from various sources of information [[11], [13]]. In addition to its role as storage, warehouse modeling is completely dedicated to the analysis of its data. The data are organized according to a multidimensional model star or snowflake [4]. These models are widely used to prepare the data for analysis. They also can produce data views commonly called data cubes. A cube of data is constituted of a set of cells, where each cell represents a fact. This is described by several axes of analysis (dimension), and is observed by one or more indicators, called measures.

The online OLAP analysis is one of the solutions for decision support. It provides tools that serve visualization, structuring, exploration data from the data warehouse based on a multidimensional representation of data which enables analyst and decision makers to process their data in an analytical, interactive, and fast way. It also allows for the ability to see the data from several dimensions. On the other hand, data mining uses learning methods to induce models of knowledge expressed in valid and understandable formalisms. Thus, we can consider that online analysis (OLAP) and data mining are two complementary disciplines. Their combination can be a potential solution to balance the weaknesses of each discipline. In addition, the structure of the data cube can provide a suitable context for applying the methods of Data Mining.

Several studies have been conducted in this direction (coupling methods of data mining with OLAP technology). In [2] they proposed a first extension of the capacities of the OLAP to describe, visualize, classify and explain. And in [[17], [10], [19], [5]], they do not share the same motivation: navigation aid and in-depth analysis of data through prediction. One difference between these proposals is given in [2]. In the context of prediction, some work in [16], is based on regression trees and consists of predicting a measurement value of new data aggregate by choosing a context analysis.

In this work, we are also interested in the prediction, but starting from partitioning a cube into sub cubes of dense data. In our case, the user selects a cell to predict construction not a context analysis. In the next section, we present an overview of existing literature as a comparative study was conducted for positioning purposes. We present in the next section, our contributions as well as our approach illustrated with an example in addition to our application in the case of real data. And in the last section, we end with a conclusion and perspectives of our work in the last section.

## 2 State of Art

Extending the capabilities of OLAP to data mining allows us to offer a wide range of advanced tools for decision support. Of course, a problem arose, related

to the size of the data, the dimensionality of the cube and hierarchies. Therefore, we present a synthesis of existing work dealing with this subject.

The adaptation of multidimensional data structure is essential for OLAP features and also or the consideration of the hierarchies of the cube in the construction of prediction model and operating results in OLAP.

Sarawagi and al. [19] use a data cube of predicted values to guide the user in operating the initial cube. Deviations between the two cubes are used to give the user three indicators. These consist of indicating, in the base cube, cells exhibiting an exceptional value as well as dimensions and cells to force to find the singular values. This approach fits perfectly in the OLAP environment.

The work of Cheng [7] aims to predict new facts. It suggests generating a new cube using a generalized linear model. The obtained cube corresponds to the prediction model.

Palpanas and al. [18] use the prediction to predict the original facts of a cube from aggregated facts. This is interesting when the user does not have all the data. In cases where the original data are available, it is suggested to specify exceptions in the data by comparing them with those approximated.

Chen and al. [5] is similar to Sarawagi and al. [19], as they both incorporate a data mining process, in fact, their approach is to identify the subset of interesting data in the light of a predictive model. The predictive model is a cube where the measurement indicates a score or a probability distribution associated with the measurement value that can be expected in the original cube. However, to predict the measurement of a non-existent fact, their effort is upstream of the construction of the prediction model and the search of the most relevant data to the learning according to the new fact that the user wishes to predict. Therefore, this proposal allows the user to optimize the accuracy of the prediction by setting the development of the model.

The proposal Y.Chen and Pei [9] is to build cubes with the results of a linear regression. A cube measuring compressible values is generated where each cube shows the general trend of the data. The measures are aggregated and provide exceptional research areas and provide data trends.

Several studies [6] deal with the same subject. From a model, the cube is synthesized to gain storage cost and also response time to queries. Most authors discuss the notion of query "*What-if* " type because the user can through the model use an estimation of the value of facts. Also, Imielinski et al. [20] used the notion of "*What-if* " at the generalization of association rules in data cubes request. [[8], [9]] then resumed this method for compressed cubes.

Finally, the work of Niemczuk [16], chosen to place more towards the needs of the user in the OLAP environment by responding more precisely. They use the learning process in a simple way and provide a prediction of the measured values as appropriate targeted user cells. Through regression trees, they provide a discrimination of the explanatory variables, and offer a prediction for empty cells.

# 3 Illustration of our approach

### Vocations

Our approach is to predict in a data cube measurement values of unexisting facts. For this reason, we propose to start by partitioning the cube into a dense subcube based on the work of [[1], [14]] which provides an analysis of potential technical probabilistic model, then in each sub data cube, we use the supervised learning method and regression tree. In our approach, the user selects a cell to predict not a context analysis as in work [16].

- Partitioning the cube of the original data (initial) into dense sub cubes according to a method of clustering,

- Building for each cubes sub dense a predictive model and validate it (machine learning and regression tree)

- Selecting the user to predict the cell rather than an analysis context,
- Predicting through the cube sub model.

### Suggested approach

To explain our approach, we use an illustration of an example of a data cube at three dimensions: Sex (f, m), Supermarket (A, F, E, B, G, I) and product (S, W, U, V, T, Y, Z). Measures correspond to the amounts of customer purchases. The data cube (*Figure 2*) is composed of 84 cells (product of cardinalities of dimensions). It is considered that out of the 84 cube cells, 12 cells are empty and whose values are to predict.

- **General notations**

We take the general notation for the structure of a data cube and a sub cube data as previously proposed in [2]. The data cube is a multidimensional representation of data (generally three dimensions, each cell of the data cube is a fact described by axes of analysis. The latter (axe of analysis) correspond to the dimension of the cube.

Definition (data cube), we consider a simple example of a three dimensions data cube (*Figure 1*): $D_1$, $D_2$ and $D_3$ . Let C be a data cube with the following properties:
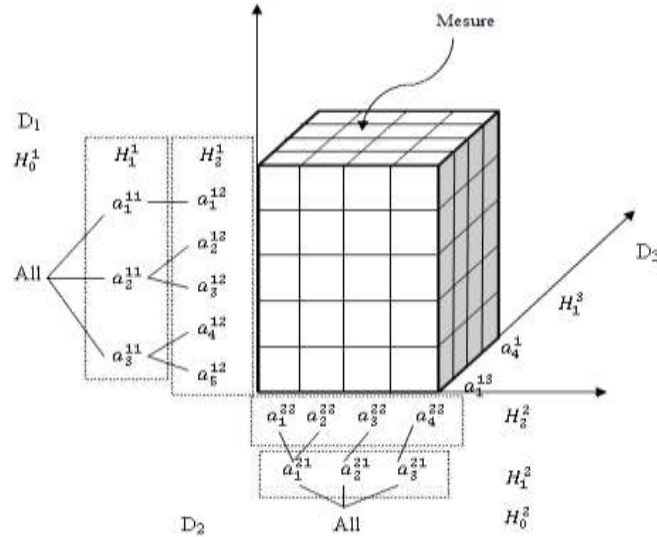


Fig. 1– Example of a data cube

- C is a non-empty set of dimensions D= {$D_i$}  $(1 \leq i \leq 3)$ ;

- C consists of a non- empty set of measurements M= {$M_q$}$(1 \leq q \leq m)$;   the measurement is the value of a cell ;

- Each dimension $D_i \in D$ contains a non-empty set of $n_i$ hierarchical levels. We consider $H_j^i$ the $j^{th}$ hierarchy level of the dimension $D_i$;

- Each hierarchical level $H_j^i \in H_i$ is a non-empty set of $l_{ij}$ modalities.  $a^{ij}$ is the $t^{th}$  modality of level; $H_j^i$ ;

- $A^{ij} = \{a_t^{ij}\}(1 \leq t \leq l_{ij})$ is the set of $H_j^i$ hierarchical levels of the dimension $D^i$ ; $A^{ij}$ is the set of modalities;

- For the total level of aggregation of a dimension, we consider that *All* is the only modality of this level. Thus, for a dimension $D_i$, we note that a $a_1^{i0} = $ *All* and $A^{i0} = \{ALL \}$ . In the following, consider a cube C with d-dimensions ($D_1$,…. , $D_i$, ….. ,$D_d$) and observed n OLAP facts according to quantitative measurement $M_q$ .

We consider a cell is full (empty respectively) if it contains a measure $M_q$ of an existing fact (or a non-existent fact).

Our proposal is to partition the data cube C in dense sub cube $C_i$ based on clustering methods. To do this we introduce the definition of a sub data cube.
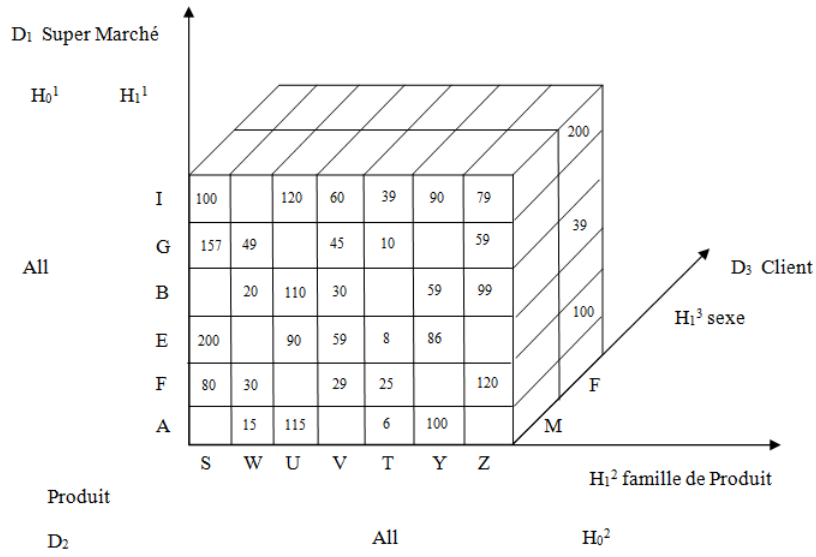


Fig. 2 – Example of a sales data cube

Definition (in the data cube) Let $D ' \subseteq D$ a non-empty set of p dimensions $\{D_1, \ldots, D_p\}$ of data cube C ($p \leq d$). The P-tuple $(\theta_1, \theta_2, \ldots, \theta_p)$ is a subset of C data along D' if and only if $\forall \ \forall i \in \{1, \ldots, p\}, \theta i \neq \emptyset$ and there exists a unique $j \geq 0$ such as $\theta_i \subseteq A_{ij}$.

Sub cube data according to a set of dimensions D' correspond to a portion of the original data cube C. It is about to set a hierarchical level $H_j^i$ in each dimension $D_i \in D$ and select in this level a sub- group $\theta_i$ which is a non-empty set of terms belonging to the set of all terms $A_{ij}$ of $H_j^i$ .

It is to be noted that a cell of a data cube C corresponds to the particular case of a sub-data cube defined by the entire set of dimensions D= $\{D_1, D_2, \ldots, D_p\}$ and as $\forall i \in \{1, \ldots, p\}$ $\theta_i$ is a singleton containing a single modality belonging to the hierarchical level of the thinnest dimension $D_i$.

### Partitioning the cube data into a dense sub-cube

The starting point of our approach is to partition the data cube into a dense sub cube with a clustering method. This idea goes back to the work of compression of cube data processed in the work of [[1], [14]] which is the approximation of the data and the search in the data cube.

Some facts to research on data approximation and mining in data cubes; very large data cubes to store and process, Data cubes are multi-way tables, high dimensional cubes with possibly useless dimensions or associations among dimensions, patterns (e.g., clusters, outliers, correlations) are hidden in large, heterogeneous and sparse data sets and users prefer approximate answers with quick response time rather than exact answers with slow execution time.

Indeed, in [14], the authors propose a Probabilistic modeling for data approximation, compression and mining in data cubes, defines focus on non-negative multi-way array factorization (NMF), and determines a potential for approximate query answering

- **Probabilistic data cube modeling**

Assume counts in cube $C=[c_{ijk}]$ arise from a probabilistic model $P(i,j,k)$. Consequently $C$ is a sample from multinomial distribution $P(i,j,k)$. Quality of Model $\theta$ is measured by the (log-) likelihood:

$$L(\theta) = \ln P(C/\theta) = \sum_{ijk} \ln P(i,j;k)$$

All models implement a trade-off between fit (high $L(\theta)$) and compression (number of parameters).

- **Non-negative multi-way array factorization**

Additive sum of **M** non-negative components:

$$P(i,j,k) = \sum_{m=1}^{M} P(m)P(i/m)\,P(j/m)\,P(k/m)$$

Each component is a product of conditionally independent multinomial distributions. So, Observations behave "the same" in each component, equivalent to decomposition of multi-way array **C**:

$$\frac{1}{N} X \approx P(i,j,k) = W_m \otimes H_m \otimes A_m$$

Such that, into non-negative factors (probabilities **W**=[*P(i,m)*], **H**=[*P(j/m)*], **A**=[*P(k/m)*])

The estimation by maximizing the log-likelihood, or equivalently the deviance defined by:

$$G^2 = 2 \sum_{ijk} c_{ijk} \ln \frac{\tilde{c}_{ijk}}{c_{ijk}}$$

Expectation-Maximization (EM) algorithm implies that iterative algorithm with multiplicative update rules and more components implies better fit, less compression.

- **Model selection: finding best trade-off**

Use Information Criteria such as AIC or BIC

$$AIC = \tilde{G}^2 - 2\,df \;\; and \;\; BIC = \tilde{G}^2 - df \; x \ln N$$

$\tilde{G}^2$ is a maximum deviance and *df* is a degrees of freedom

- **Rates of compression and approximation:**

For approximation measured by deviance $G^2$: $G^2$=0 means perfect approximation (saturated model). Higher $G^2$, worse approximation

For compression: How much smaller is the model? Compression rate it consists in ratio of parameters over cells:

$$R_t = 1 - \frac{f}{N_t} = \frac{df}{N_t}$$

Where, *df* degrees of freedom and $N_t$ is number of cells

For NMF:    $R_t = 1 - M \frac{I+J+K-2}{IJK}$ , such as, *M* is number of components

- **Approximate query answering:**

Query reformulation on NMF components select a portion of the cube (*Slice* and *Dice* differ on the extent of the selection). Probabilistic model cuts the processing time as; only necessary cells need to be calculated (no need to compute entire cube) and irrelevant (i.e., outside of the query scope) components may be ignored. Saving is important if query selects a small part of the cube and components are well distributed.

In our example of three dimensional data cube we design a slice of the cube.

-Approximate query answering: *Roll-up:*

-Aggregate values over all (or subset of) modalities of one or several dimensions

-Easily implemented by summing over probabilistic profiles in the model

For example, roll-up over dimension k:

$$\sum_{k=1}^{K} P(i,j,k) = \sum_{m=1}^{M} P(m)P(i/m)\,P(j/m) \quad \sum_{k=1}^{K} P\left(\frac{k}{m}\right) = \sum_{m=1}^{M} P(m)P(i/m)\,P(j/m)$$

Where ;

$$P(i,j,k) \approx \frac{C_{ijk}}{N} \quad and \quad \sum_{k=1}^{K} P\left(\frac{k}{m}\right) = 1$$

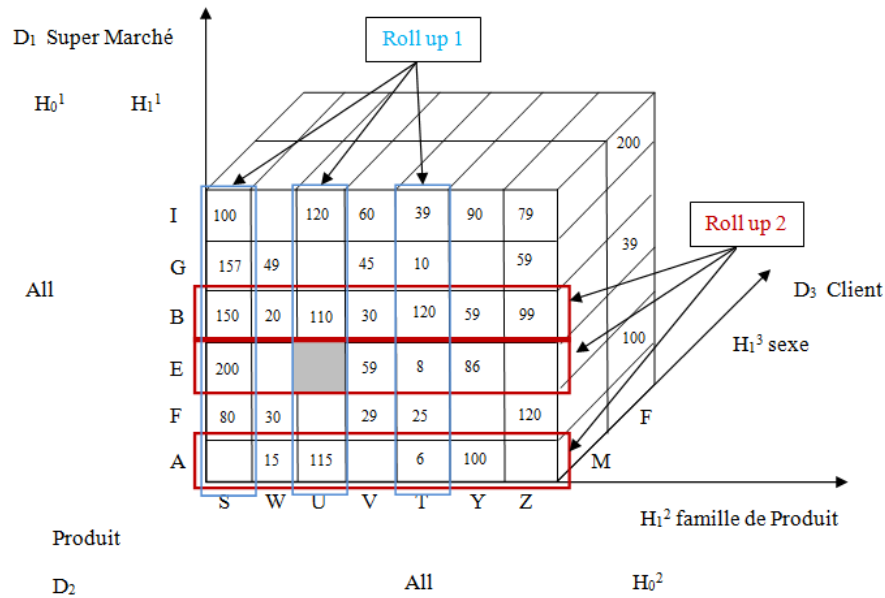Get rolled-up model "for free" from original model, Roll-up on model much faster than on data



Fig. 3- Example of resulting sub cube dense

### Construction and validation of the predictive model:

Sub data cube dense so designated, may be considered as regions (rich) carrying information and can be considered as a training set to build the prediction model.

Several methods proposed in the literature, suggest the construction of a regression tree such as, the CHAID method [12], AID [15], CART [3] and recently Arbogodaï [21]. We use one of its methods (a comparative study is intended).

CART allows the creation of binary tree based on the vocabulary of supervised learning methods, the explanatory variables are the dimensions of sub cubes, and the variable to predict represents corresponding $M_q$ measure.

Each dense sub cube $C_i$ is partitioned into a learning sample and a test sample, the model thus constructed will be based on 70% of the facts used for learning and the rest 30% of the facts for the assessment of the model. For each sub cube, the construction of the regression tree is made by recursive partitioning; in fact, algorithm tries to predict the dependent variable from the best predictor. For this fact, the variables are grouped into two or more sub-sets.

The associated sub cube data prediction is the average of observations belonging to the grouping and consistency of a group is measured by the variance of the variable to predict in the peak, when segmenting a peak in two or more sub overall, it seeks to minimize the intra-group variance and maximize inter-group variance.

The criteria for evaluation of a regression tree are average; which is the difference between the true value of the variable to be predicted and the observed value, and error reduction; which measures the proportion of variance explained in the model i.e. the quality of the regression. However, if the error rate is close to 0, then the model is not incorrect and also if the reduction of error is equal to 0, then we have a perfect prediction.

The model is validated if the average error and error reduction in test phase are low and similar to that obtained in the learning phase. For our example, the average error in the first cube is 0.142 which acceptable.

- **Interpretation**

For each dense sub cube $C_i$ correspond a regression tree, which makes k a decision rule. The rules of a model are denoted $R_i = \{ R_i^1, R_i^2, \ldots, R_i^k )$

Definition (Rule decision) Let R (X $\Rightarrow$ Y, S, $\sigma$) a decision rule, the predicate X is the conjunction and / or disjunction of terms corresponding to a history of the rule. Y is the average predicted value for measurement $M_q$ given X. S is the number of individuals satisfying X and $\sigma$ is the standard deviation of $M_q$ overall sample checking X indicating the homogeneity of the facts supporting the rule.

In our example, we have four regression trees, each of which returns a set of decision rules, for example in the first dense sub cube, we have the following rules:

$R_1^1((A \vee B)) \Rightarrow (117{,}821; 66{,}7\% ; 0{,}24)$

$R_1^2(M \wedge E) \Rightarrow (101{,}214 ; 33{,}3\%; 0{,}54)$

$R_1^3((F \wedge (A \vee E)) \Rightarrow (128{,}592 ; 33{,}3\%; 0{,}35)$

$R_1^4(M \wedge (A \vee B) \Rightarrow (107{,}048 ; 33{,}3\%; 0{,}94)$

For example, Rule 2 shows that if the customer is male and the market is E, the profile will be 101.214, 33.3% of customers belong to this category and the standard deviation is 0.54.

### Prediction of measuring a cell:

Unlike the approach of [16], where the user selected a context analysis, in our approach, the user designates directly the cell to predict and to determine within which sub cube the cell belongs and directly predict the measurement of the cell through predictive model built on the cube.

The user selects an empty cell to predict (with no value for measuring $M_q$ (c) = Null):

- Let c be the empty cell to predict selected by the user,

- Let us seek the sub data cube $C_i$ as C belongs to $C_i$ ,

- $M_q$(c) denotes the value of measurement taken by the cell $M_q$ ,

- Predict based on the predictive model of the sub cube: We are looking for the rule $R_i$ derived from the regression prediction model built from the sub cube $C_i$ as its antecedent X for all terms describing c. For a given rule, we only look at conjunctions of its terms.

**Let's proceed as follows:**

Algorithm:

---

*For $M_q$ (c) = NULL do*
　　*if  c ∈ $C_i$, then*
　　　*For $R_i^k$ ∈ $R_i$, do*
　　　　　　　　*$M_q(c)$  ←  Y*
*End for*
　　*Else*
　　　*For All dense cube*
　　　　　*For all "DRPM" (decision rules of the predictive model)*
　　　　　　　*If X has modalities describing "c" ( X is antecedent of DRPM)*
　　　　　　　*And having a large effective and a minimal error then:*
　　　　　　　　*$M_q(c)$  ←  Y*
　　　　　　*End if*
　　　　　*End for*
　　　　*End for*
　　　*End if*
*End for*

---

The user designates the cell c to predict in the data cube, we determine the dense sub cube to which the cell c belongs. Note that $M_q$ (c) the measured value $M_q$ which takes the cell c. for each empty cell selected by the user as $M_q$ (c) = *Null*. We find the dense cube $C_i$ to which the cell c belongs. Then, we search for the decision rule $R_i$ as its antecedent X has all its terms describing c. Thereafter we assign to the measurement of cell $M_q$ (c) the average Y value of the prediction rule.

In our illustrative example, if the user selects the gray cell (Figure 3) described by the terms (M, E, U) for the dimensions, Sex, Super market and product user, we will look for the sub cube to which the cell belongs and using the predictive model, the rule $R_1^2(M \wedge E) \Rightarrow (101,214\,; 33,3\%; 0,54)$ be selected. Therefore, male clients will have a purchasing average of 101,214 at supermarket E.

Thus, for each empty cell selected by the user, we affect the value of the decision rule derived from the dense sub cube to which it belongs.

If c doesn't belong to any dense sub cube, we look for all decision rules of the predictive model; such as its antecedent X has a modalities describing c and having a large effective and a minimal error, we assign to $M_q(c)$, the value Y.

The choice of use of dense sub cubes is justified by; on the one hand, the quality of information carried by these dense sub cubes, on the other hand, when the designation of the cell to predict by the user, it will be interesting to search in

the predictive model dense sub cube where the cell belongs than look through of all data cube.

## 4 Conclusion

In the continuity of the work treating the coupling between the on-line analysis and data mining, we have made an approach to prediction from dense sub cube of a data cube.

We tried to give a summary of the work developed in the prediction in OLAP. Chen et al. [5] suggested a prediction model in the form of a cube, left to the user for operation, while in [16], the selected user context analysis to estimate the value of the measurement of non-existent facts. In our approach, the user selected an empty cell to predict due to the predictive model in dense data cube to which it belongs.

We wish to develop a comparative study of regression tree methods, as we also discuss the very important issue of association rules. In this context, our method would be enriched in that the user has the possibility to choose between different types of rules: simple rules and more complex ones. These include intervals or combinations of terms of their antecedents and consequences. Finally, we deal with our method with existing work in qualitative variable context. The main objective will be looking for methods that are not too far from our context. And then, if possible, pick one and explore the possibility to find a connection between them.

## References

[1] D. Barbara and X. Wu, Using log linear models to compress data cubes, In WAIM'00: Proceedings of the First International Conference on Web-Age Information Management, pages 311–322, London, UK, 2000. Springer-Verlag.

[2] R. Ben Messaoud, Couplage de l'analyse en ligne et de la fouille de données pour l'exploration, l'agrégation et l'explication des données complexes. PhD thesis, Université Lumière Lyon 2, Lyon, France, Novembre 2006.

[3]   L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. Classification and Regression Trees. 1984.

[4]   S. Chaudhuri, Dayal, An Overview of Data Warehousing and OLAP Technology, SIGMOD Record, 26(1): 65–74. 1997.

[5]   B. –C. Chen, L. Chen, Y.Lin, and R. Ramakrishnan. Prediction Cubes. Proceeding of the 31th International Conference on Very Large Data Bases (VLDB'05), pages 982-993, Trondheim, Norway, August – September 2005. ACM Press.

[6]   B. –C. Chen, R. Ramakrishnan, J.W. Shavlik, and P. Tamma. Bellwether Analysis: Predicting Global Aggregates from Local Region. Proceedings of the 32th International Conference on Very Large Data Bases (VLDB'06), pages 655-666, Seoul, Korea, September 2006. ACM Press.

[7]   BS. Cheng. Statistical Approaches to Predictive Modeling in Larage Data-bases. Master's thesis, Simon Fraser University, British Columbia, Canada, February 1998.

[8]   G. Dong, J.Han, J.M.W. Lam, J. Pei, and K. Wang. Mining muti-dimensional constrained gradients in data cube. In VLDB'01: Proceedings of the 27$^{th}$ International Conference on Very Large Data Bases, pages 321-330, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[9]   J. Han, J. Wang, G. Dong, J. Pei, and K. Wang. Cube explorer: online exploration of data cubes. SIGMOD'02: Proceedings of the 2002 ACM SIG'02 international conference on Management of data, pages 626-626, New York, NY, USA, 2002. ACM.

[10]  T. Imielinski, L. Khachiyan, and A. Abdulghani. Cube grades (2002): Generalizing association rules. Data Mining and Knowledge Discovery, pages 219-257.

[11]  W.H. Inmon, Building the Data Warehouse, John Wiley & Sons. 1996.

[12]  G. V. Kass. An exploratory technique for investigating large quantities of categorical data. Applied Statistics, 29(2): pages: 119-127, 1980.

[13]  Kimball R., The Data Warehouse Toolkit, John Wiley & Sons. 1996.

[14]  R. Missaoui, C. Goutte, A.K. Choupo, B A. oujenoui, A Probabilistic Model for Data Cube Compression and Query Approximation, Proceedings of the

10th ACM International Workshop on Data Warehousing and OLAP (DOLAP'2007), pages 33–40, Lisbon, Portugal : ACM Press. November 2007.

[15] J. N. Morgan and J. A. Sonquist. Problems in the analysis of survey data, and a proposal. Journal of the American Statistical Association, 58(302): page: 415-434, 1963.

[16] A B. Niemczuk. Vers l'intégration de la prédiction dans les cubes OLAP. Master's thesis. Université Lumière Lyon 2, Lyon, France, Juin 2007.

[17] T. Palpanas and N. Koudas. Entropy Based Approximate Querying and Exploration of Datacubes. July 2001.

[18] T. Palpanas, N. Koudas, and A. Mendelzon. Using Data cube Aggregates for Approximate querying and Deviation Detection. IEEE Transactions on Knowledge and Data Engineering, 17 (11): 1465-1477, Novembre 2005.

[19] S. Sarawagi, R. Agrawal, and N. Megiddo. Discovery-driven Exploration of OLAP Data Cubes. In Proceeding of the 6th International Conference on Extending Database Technology (EDBT'98), pages 168-182, Valencia, Spain, Mars 1998. Springer.

[20] T. Palpanas, N. Koudas and A. Mendelzon. Using Data cube Aggregates for Approximate Querying and Deviation Detection. IEEE transactions on knowledge and Data Engineering, 17 (11): 1465-1477, November 2005.

[21] D. A. Zighed, G. Ritschard, W. Erray, and V. M. Scuturici. Abogodaï, a new approach for decision trees. In 7th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 03), Dubrovnik, Croatia, volume 2838 of LNAI, pages 495-506, Heidelberg, Germany, September 2003. Springer.