

Publicação: Fevereiro de 2019

Digital Object Identifier https://doi.org/10.33911/singular-etg.v1i1.18

Market Basket Analysis in a Financial Institution

FELIPE DE MOURA REZENDE DOS SANTOS, MARCELO LADEIRA.

Computer Science Department, University of Brasília-(UnB), P.O. Box 4466, 70910-900, Brasília-DF, Brazil. (e-mail:feliperezende86@gmail.com, mladeira@unb.br)
Autor Correspondente: Felipe Rezende (e-mail: feliperezende86@gmail.com).

ABSTRACT - This article demonstrates a study on Market Basket Analysis of a financial institution, showing rules of personal consumer association of the state of São Paulo. A concept about three association algorithms is presented, but a study with only one is performed. The paper is divided into an introduction, describing a brief account of the reason for choosing the subject. Understanding the business, where it is explained about the financial institution and the importance of the study to the institution. The way the data are handled is demonstrated in Understanding the Data, just as the Data Preparation is described in the sequence, putting all the filters and treatments that were done on the data. In the following, it is described the Modeling, which reports on algorithms of association rules and on examples of these algorithms, as well as which algorithm was chosen to be treated in the paper. Evaluation explains on the results obtained with the study and the Implementation as it was done all the analysis of the data and the results obtained. Finally, we have the Conclusion about the learning obtained with the article and what future work to do.

KEYWORDS - Data mining, Data, Mining, Association, Apriori

I. INTRODUCTION

What does beer have to do with diapers?

It was a common question until one of the largest US retailers conducted a study and found that diaper sales were associated with beer. In general, buyers were men who went out at night to buy diapers and ended up buying some beer. The products were placed side by side and the result was increased sales of the two products. It is unknown whether this study actually occurred or whether it is only a "legend" about data mining study.

The competitiveness of companies is becoming increasingly dependent on the quality of their decision-making. Therefore, it is no wonder that companies often try to learn from past transactions and decisions in order to improve the quality of decisions made in the present or future. To support this process, large amounts of data are collected and stored during business operations. Later, this data is analyzed for relevant information. This process is called data mining or knowledge discovery in databases. Data mining is relevant to many different types of companies. As examples, retail stores obtain customer profiles and their purchasing patterns and supermarkets analyze their sales and the effect of advertising on sales. This "targeted marketing" is becoming increasingly important.

Thus, this case study tends to propose a specific group of products purchased by customers of a financial institution and predispose a tying or otherwise known as cross-selling, ie, point where the product this customer is likely to purchase, by consumption profile.

II. UNDERSTANDING THE BUSINESS

Every financial institution has various types of products to be offered to customers, credit products (loans), investment products (investment funds) and services products (insurance). In some cases they are offered by the attendants, products that are not useful for the client profile, generating a certain discomfort and a problem in the accomplishment of this purchase.

It will be used Data Mining concepts to have a predisposition of certain items to be offered in conjunction with products that the customer already has, increasing the probability of making this offer effective. Through association of the products already owned by him. For this, all the products contracted by the customer will be determined, so that a grouping of this information can be made and a profile of customers and products acquired be generated. This process would have the cost of being processed in a large database, which is not prepared for such a procedure, running the risk

of not being able to perform in a timely manner to carry out the analysis. Therefore a study on a sample of data was determined. Only one month, individual customers and products used.

As a criterion of success, one can place the possible creation of customer profile groups with the same interest of consumption of the products of the Financial Institution. With this having the possibility of a better offer of consumption of other products for clients who are doing business with the institution.

III. UNDERSTANDING DATA

We have a lot of customer data. In order to start this project, it was defined that a study of only 1 month, in the case, July 2017, would be done because it is the most recent month already processed. Nevertheless, it remained with more than 40 million lines, with that restricted to only individuals that contain current account in the institution. For a more restricted analysis of data, the state of Sao Paulo was chosen, with a total~ of 9,343,551 lines.

The database used to conduct this study contains 160 variables and more than 1 billion lines, since it contains data referring to all operations of all clients of the financial institution since January 2011. This database is partitioned by year and is indexed by year and month.

IV. DATA PREPARATION

The database used in the study contains no problems with the need to be corrected to be a basis of use for disclosure of results of products for the dependencies, so it was only required to make necessary filters and the reduction of variables needed to study.

Initially, a filter was performed to have only the data of July 2017, excluding operations that do not contain customers, bringing only the customer code, the unit of the federation of cadastre and the products contracted by him in the institution. Altogether, more than 236 million lines were generated, with 1 line per customer and product of this customer.

After it was decided that it would be restricted to a specific state for the purpose of studying and executing the algorithms in a home computer, the state of São Paulo was chosen. Included in the preparation of the data were the permanence of only the Individuals customers. Totaling approximately 9 million lines.

A pivoting of the data was performed, generating a column for each product of the bank, having for each customer all the existing products in their relation with the bank.

V. MODELING

The analysis of purchases and product records typically uses association rules representing the relationship patterns found between the data items of the analyzed set. In databases where records are transactions, these rules are known as transactional association rules, whereas in the case of databases where records represent clients, accounts, products, services

and others, they involve multiple attributes and, therefore, are called multidimensional association rules.

In the context of marketed product sales, or cross-selling the membership rules allow a store to recommend product B for the customer who is buying product A, since it knows the rule, for example, that 30% of your customers who purchase A also buy B. Thus, the customer is encouraged to buy more products, which may possibly interest him, based on consumption characteristics of previous purchases in the system. In this way, not only the quantity of sales is maximized, but also the quantity of sales of a particular product.

Association rules were introduced [1] as follows: are I = i1,i2,..., in a set of m distinct items and D a database formed by a set of transactions, where each transaction T is composed of a set of items such that $T \subseteq I$. An association rule is an expression of type XY where $X \subseteq I$, $Y \subseteq I$, $X \in \varnothing$, $Y \in \varnothing$ and $X \cap Y = \varnothing$, that is, both the antecedent (X) and the consequent (Y) of an association rule can be formed by sets containing one or more items, but can not be empty and are independent sets since they have no items in common. For example:

{Cerveja, Faldas} → {Leite} {Cerveja, Leite} → {Fraldas} {Fraldas, Leite} → {Cerveja} {Cerveja} → {Fraldas, Leite} {Leite} → {Cerveja, Fraldas} {Fraldas} → {Cerveja, Leite}

Figura 1. Exemple of Association Rules

The importance of an association rule can be measured in terms of support and confidence. The support of a rule determines how often a rule is applied to a dataset, that is, the probability that the first term of the implication is true. The confidence of a rule determines how often the items in Y appear in transactions that contain X. In the previous example, the likelihood of a customer buying beer and diapers (ie, the term X is true) is referred to as support, while the conditional probability of a customer buying milk, since he bought beer and diapers, is referred to as confidence.

The problem of association mining is to find all association rules that have support and trust greater or equal, respectively, to a minimum support (SupMin) and a minimum trust (ConfMin), specified by the user.

In fact, for a long time, the search for association rules was of exclusive interest to applications that deal with information from Market Baskets, which led to this technique being commonly called market basket analysis (MBA).

However, we can see that association rules can be extracted from any database where there are implicit relationships



between the different attributes. The usefulness of these rules is not only in the generation of new knowledge, but also in the confirmation of business rules that are used but never proven. Thus, association rules apply to several business areas such as studies of access to computers, search for new clients, population census and medical information analysis, among others.

The discovery of standards through analysis of association rules is an important one to support decision making. With this information, a manager can identify new business opportunities, better know their customers and their buying profiles, identify products that influence the sale of others, and various other information that can increase their competitiveness in the market.

The first algorithms to be used in the discovery of association rules were the artificial immune system (AIS) [2] and set-oriented mining (SETM) [3]. In most of the more recent studies the Apriori algorithm has been widely used, as well as its variations. In this work, due to being used in more recent studies and because of the shortage of time for further deepening of other algorithms, Apriori was chosen and used.

A. ARTIFICIAL IMMUNE SYSTEM

The technique used by the artificial immune system (AIS) generates and counts sets of items as you read the database transactions.

For each transaction, it is determined which of the largest sets found in the previous transaction are also in the current transaction and new sets of items are generated by extending these sets with items from this record.

The disadvantage is that this method generates and counts unnecessarily sets of items that are considered small. The following figure illustrates how it works.

The first pass of the algorithm in the database makes the search for baskets of only one item counting how many times these are repeated in the transactions, ignoring those that occur only once. The second pass in the base searches for baskets with two items, counting how many times they appear together in the scanned records.

In this example, the last pass contains baskets of three items, which are the largest basket that is often larger than one. Thus, the algorithm has as solution the set 2,3,5.

B. SET-ORIENTED MINING

Like AIS, SET-oriented Mining (SETM) generates the candidate item sets at the time it is going through the database transactions. However, counting is only performed at the end of the sweep of the base.

The transaction identifier is stored together with the candidate set in a sequential structure.

At the end of the reading of all records, the most frequent set of items is determined by the aggregation of the sequence structure.

In addition to having the same disadvantage as AIS, for each candidate set, there are several calculated size values.

The following figure illustrates the steps in the algorithm.

Base de Dados	
ID	Itens
100	1, 3, 4
200	2, 3, 5
300	1, 2, 3, 5
400	2,5

1a Passagem C1	
Conjunto Candidato	Contagem
{1}	2
{2}	3
{3}	3.
{5}	3

2a Passagem	· ·
Conjunto Candidato	Contagem
{1,3}*	2
(1,4)	1
{3,4}	1
(2, 3)*	2
{2,5}*	3
(3, 5)₹	2
{1,2}	1
{1.5}	1

*Largest set of items in iteration i

3a Passagem C3	
Conjunto Candidato	Contagem
{1,3,4}	1
{2,3,5}*	2
(1, 3, 5)	1

Figura 2. Example of Artificial Immune System

In SETM, the first pass of the database results in the frequency of occurrence of each item in the transactions, as in the previous example based on AIS. However, the second sweep in the base results in a list of combinations of product pairs where each combination receives the identifier of the transaction from which it was extracted.

Thus, we will have repeated combinations with different identifiers such as the basket 1,3 that in this example occurs in transaction 100 and in transaction 300.

The last pass at the bottom reveals the baskets of three items and the solution found is the one that appears in most transactions.

With this, we also have the set of items 2,3,5 as a result of applying SETM to this recordset.

C. APRIORI ALGORITHM

This algorithm was the first tool to discover association rules in databases with large volumes. Several modifications were



Base de Dados		
ID	ID Itens	
100	1, 3, 4	
200	2, 3, 5	
300	1, 2, 3, 5	
400	2, 5	

1a Passagem C1	
Conjunto Candidato	Contagem
{1}	2
{2}	3
{3}	3
{5}	3

2a Passagem C3	
Conjunto Candidato	ID
(1, 3)	100
(1, 4)	100
(3, 4)	100
(2, 3)	200
[2, 5]	200
(3, 5)	200
(1, 2)	300
(1, 3)	300
(1, 5)	300
(2, 3)	300
(2, 5)	300
(3, 5)	300
(2, 5)	400

3a Passagem C3	
ID	
100	
200	
300	
300	
֡	

Figura 3. Example of SET-oriented Mining

proposed to improve their efficiency and similar algorithms were proposed [4] [5], introducing more expressive rules.

First, the algorithm identifies the items that are part of each transaction. The algorithm then determines the membership rules to enter these items by selecting the associations that occur most frequently (or larger) in the set of transactions in question.

The largest set of candidate items from the previous step is

Base de Dados		
ID	Itens	
100	1,3,4	
200	2,3.5	
300	1,2,3,5	
400	2,5	

1a Passagem C1	
Conjunto Candidato	Contagem
{1}	2
{2}	3
{3}	3
(5)	3

Conjunto Candidato	Contagem
{1, 2}	2
{1,3}*	1
{1,5}	1
(2, 3)*	2
(2.5)*	3
{3.5}*	2

(1, 2, 3)	
{1, 3, 5}	
(2, 3, 5)	

3a Passagem	C3
Conjunto Candidato	Contagem
{2, 3, 5}*	2

Figura 4. Exemple of Apriori

taken to the next iteration, generating other sets larger than 1.

Finally, the algorithm eliminates the generated sets that have a subset that is not the largest. Based on the most frequent set of items, membership rules that meet minimum support and confidence values are generated. As a result, all sets of common items are discovered, producing all association rules that respect those limits.

As in the previous examples, the first pass of the Apriori algorithm in the database results in the listing of the frequency of the items in the transactions (baskets of an item only). In the next step, the generated list contains the baskets of two items with their respective frequencies.

Of these baskets only the most frequent baskets will be considered and, thus, the baskets with three items are already prepared for the last step, which will only then count the baskets more frequent related in the previous step. With this, the result is the basket of items 2,3,5.

Número 1, 2019 9



VI. EVALUATION

Based on the results obtained, we can successfully evaluate the study because we were able to respond to the initial questioning, determining a shopping basket of products of the financial institution under study.

The process has been reviewed and tested countless times, changing conditions and rules to produce the necessary information demonstrated in the next chapter.

As next steps we can list the study of other algorithms for the association of shopping baskets. As well as the inclusion of other ways to predict possible purchases from certain groups of customers.

VII. IMPLEMENTATION

All customer information base was obtained through SAS at the financial institution, where the month of July and the year 2017 were filtered, as well as restricting only individuals from the state of São Paulo who have a bank account with the financial institution.

This data is confidential inside the financial institution, because of this, we can not disponibilize best explications or the real data with this study.

In view of the large volume of data without these filters, a total of more than 35 million lines were still present. After pivoting, considering each customer's product as a specific column, it dropped to 9,343,551 million lines. Where it was divided into females and males with a total of 3,879,539 and 5,464,012 lines respectively. The data was exported in CSV format for use and study of the mined information.

Analyzing women's data, we have the most common items in addition to the account (which was the filter criterion), 100% have a savings account and 82.21% have a credit card in their product basket, 47% service via cellular SMS and 36.6% has service package. We can consider these 4 items as the basic basket of products for women and from there check the items to be offered for new hires.

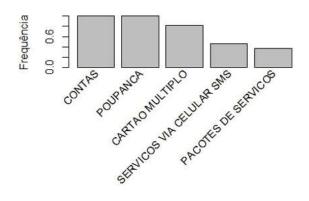


Figura 5. Frequency of women's products

By performing the Apriori algorithm by RStudio, with a support of 5% and 90% confidence we found 580 rules where ordered by lift and presenting only the first 5 rules, would be as follows: In this case, we can verify that, with a high

	1hs		rhs			support	confidence	lift	count
[1]	{CARTAO MULTIPLO, PREVIDENCIA ABERTA} =	*>	{FUNDOS	DE	INVESTIMENTOS}	0.053	0.97	10	204116
[2]	{CARTAO MULTIPLO, POUPANCA,								
	PREVIDENCIA ABERTA} =	*>	{FUNDOS	DE	INVESTIMENTOS}	0.053	0.97	10	204116
[3]	{CARTAO MULTIPLO, CONTAS, PREVIDENCIA ABERTA} =	-	(FUNDOS	D.F.	TABLESTMENTOS	0.053	0.97	10	204116
[4]	{CARTAO MULTIPLO, CONTAS, POUPANCA,					0.033	0.97	10	204110
	PREVIDENCIA ABERTA} =	=>	{FUNDOS	DE	INVESTIMENTOS}	0.053	0.97	10	204116
[5]	{PREVIDENCIA ABERTA} =	=>	{FUNDOS	DE	INVESTIMENTOS}	0.055	0.96	10	212600

Figura 6. Random Rules for Female Customers

confidence and lift, we have, in all cases, when a person has Open Retirement, they tends to have Investment Funds.

Another form of research would be to assist in the goals set for agencies. Where we can check people who own a particular product, which products already exist in your bank basket. Determining the audience to be offered this product. As an example, we use the Capitalization Title with a single payment, in this case, we use as the sorting criterion the confidence that indicates the probability that a transaction that has an item lhs also contains the item rhs. And we kept showing the first 5 items, in order to find 160 rules.

[1]	Ths {CARTAO MULTIPLO, CHEQUE ESPECIAL, EMISSAO DE TITULOS,	rhs	support	confidence	lift	count
	PAGAMENTO MENSAL) =>	TITULOS DE CAPITALIZA PAGAMENTO UNI		0.45	11	4322
[2]	(CARTAD MULTIPLO, CHEQUE ESPECIAL, EMISSAO DE TITULOS, POUPANCA, TITULOS DE CAPITALIZACA PAGAMENTO MENSAL) =>	(TITULOS DE CAPITALIZA	ACAO ICO} 0.0011	0.45	11	4322
[3]	{CARTAO MULTIPLO, CHEQUE ESPECIAL, CONTAS, EMISSAO DE TITULOS, TITULOS DE CAPITALIZAC/ PAGAMENTO MENSAL} *>	(TITULOS DE CAPITALIZA	NCAO ICO} 0.0011	0.45	11	4322
[4]	[CARTAO MULTIPLO, CHEQUE ESPECIAL, CONTAS, EMISSAO DE TITULOS, POUPANCA, TITULOS DE CAPITALIZAC, PAGAMENTO MENSAL] =	TITULOS DE CAPITALIZ	ACAO NICO} 0.0011	0.45	11	4322
[5]	{CHEQUE ESPECIAL, EMISSAO DE TITULOS, TITULOS DE CAPITALIZAC, PAGAMENTO MENSAL} =	- {TITULOS DE CAPITALIZ	ACAO urco) 0.0011	0.45		

Figura 7. Target specific for women clients

When the agency receives a task of executing Capitalization Bonds with a single payment, we can consider, with confidence of 45%, the following situations of customer baskets:

- 1) Multiple Card, Special Check, Issuance of Securities, and Capitalization Bond with monthly payment;
- Multiple Card, Special Check, Issuance of Securities, Savings and Capitalization Bond with monthly payment;
- Multiple Card, Special Check, Issuance of Securities, Accounts and Capitalization Bond with monthly payment;
- 4) Multiple Card, Special Check, Issuance of Securities, Savings, Accounts and Capitalization Bond with



monthly payment;

5) Multiple Card, Issue of Securities and Capitalization Bond with monthly payment;

In all cases, we had a support of 0.11% and a lift of 11. More than 4,300 rules were found in each of the baskets. When we analyzed male clients, we found the most common items were the multiple card with 79.9%, Cellular SMS Services with 44.3%, Savings with 36.2% and Services Package with 32.2In the case of the algorithm Apriori, with a support of 5% and confidence of 60%, we found 580 rules where ordered by lift and presenting only the first 5 rules, would be as follows:

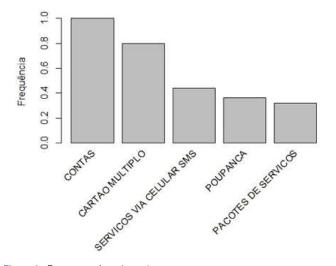


Figura 8. Frequency of men's products

	1hs		rhs		support	confidence	lift	count
[1]	{CARTAO MULTIPLO, PACOTES DE SERVICOS,							
	SEGURO DE PESSOAS}	=>	{CDC	EMPRESTIMO}	0.054	0.66	5.0	295120
[2]	{CARTAO MULTIPLO, CONTAS,							
	PACOTES DE SERVICOS,							
12353	SEGURO DE PESSOAS}	=>	{CDC	EMPRESTIMO}	0.054	0.66	5.0	295120
[3]	{PACOTES DE SERVICOS,							
	SEGURO DE PESSOAS}	=>	{CDC	EMPRESTIMO}	0.055	0.66	5.0	300510
[4]	{CONTAS, PACOTES DE SERVICOS,							
	SEGURO DE PESSOAS}	=>	{CDC	EMPRESTIMO}	0.055	0.66	5.0	300510
[5]			(
	SERVICOS VIA CELULAR SMS	} =>	{CDC	EMPRESTIMO}	0.065	0.60	4.6	357393

Figura 9. Random Rules for Male Customers

In this case, we can verify that we have, in all cases, when a person has People Insurance they tends to have CDC Loan. Differentiating from the goal set for women, let's put it that we want to hire Auto Insurance for men, what baskets of customers should we prioritize for contacts? The first 5 items were kept, sorted by confidence. Found 160 rules.

When receiving a task to perform Auto Insurance for men, we can consider, with confidence of 39% the following situations of customer baskets: 1 - Multiple Card, Special Check, Investment Fund, Receipt of Securities, Asset Insurance, Personal Insurance, Cellular SMS Service; 2 - Multiple Card, Special Check, Accounts, Investment Fund, Receipt of Securities, Asset Insurance, Personal Insurance, Cellular

[1]	Ths {CARTAO MULTIPLO, CHEQUE ESPECIAL, EMISSAO DE TITULOS, TITULOS DE CAPITALIZACAO	rhs		support	confidence	lift	count
	PAGAMENTO MENSAL) =>	{TITULOS	DE CAPITALIZAÇÃO PAGAMENTO UNICO}	0.0011	0,45	11	4322
[2]	{CARTAO MULTIPLO, CHEQUE ESPECIAL, EMISSAO DE TITULOS, POUPANCA, TITULOS DE CAPITALIZACAO						
	PAGAMENTO MENSAL} =>	{TITULOS	DE CAPITALIZAÇÃO PAGAMENTO UNICO}	0.0011	0.45	11	4322
[3]	{CARTAO MULTIPLO, CHEQUE ESPECIAL, CONTAS, EMISSAO DE TITULOS, TITULOS DE CAPITALIZACAO PAGAMENTO MENSAL} =>	{TITULOS	DE CAPITALIZAÇÃO				
[4]	(CARTAO MULTIPLO, CMEQUE ESPECIAL, CONTAS, EMISSAO DE TITULOS, POUPANCA, TITULOS DE CAPITALIZACAO		PAGAMENTO UNICO}	0.0011	0.45	11	4322
	PAGAMENTO MENSAL} =>	{TITULOS	DE CAPITALIZACAO PAGAMENTO UNICO}	0.0011	0.45	11	4322
[5]	{CHEQUE ESPECIAL, EMISSAO DE TITULOS, TITULOS DE CAPITALIZAÇÃO						
	PAGAMENTO MENSAL} =>	{TITULOS	DE CAPITALIZAÇÃO PAGAMENTO UNICO)	0.0011	0.45	-11	4348

Figura 10. Target specific for men clients

SMS Service; 3 - Special Check, Investment Fund, Securities Receipt, Asset Insurance, Personal Insurance, Cellular SMS Service; 4 - Special Check, Accounts, Investment Fund, Receipt of Securities, Asset Insurance, Personal Insurance, Cellular SMS Service; 5 - Multiple Card, Special Check, Investment Fund, Receipt of Securities, Asset Insurance, Personal Insurance; In all cases, we had a support of 0.11% and a lift of 18. More than 6,000 rules were found in each of the baskets.

VIII. CONCLUSION AND FUTURE WORK

Three association rule algorithms were presented for Market Basket Analysis. It was conceived on the three algorithms, in the case AIS, SETM and Apriori, it was verified that both work in a very similar form and due to other recent studies prefer Apriori, being thus, the algorithm chosen for the study of the case.

Based on customers from the state of São Paulo, the most common products of each sex and which groups are most common in both cases. In this way it was verified that the interest between men and women differ in relation to the preference of purchase of products of the institution. It was also verified, given a product as "goal" of increment in the baskets of the customers, for which profile of customers should be offered these products. And we were able to achieve the goal proposed by the study, showing customers that have certain product groups, tend to acquire the product requested in the campaign, and thus, will have a greater probability of success in the offer to this profile of customers.

Referências

- R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules' between sets of items in large databases," in Acm sigmod record, vol. 22, pp. 207–216, ACM, 1993.
- [2] D. Dasgupta, Z. Ji, and F. Gonzalez, "Artificial immune system (ais) research in the last five years," in Evolutionary Computation, 2003. CEC'03. The 2003 Congress on, vol. 1, pp. 123–130, IEEE, 2003.

Número 1, 2019 11



- [3] M. Houtsma and A. Swami, "Set-oriented mining for association rules in relational databases," in Proceedings of the Eleventh International Conference on Data Engineering, pp. 25-33, Mar 1995.
- S. Brin, R. Motwani, J. D. Ullman, and S. Tsur, "Dynamic itemset counting and implication rules for market basket data," in ACM SIGMOD Record, vol. 26, pp. 255–264, ACM, 1997.

 [5] J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate
- generation," in ACM sigmod record, vol. 29, pp. 1-12, ACM, 2000.



FELIPE DE MOURA REZENDE DOS SAN-TOS Possui graduação em Sistemas de Informação pela Faculdade de Pindamonhangaba (2008), MBA em Engenharia de Software (2014) e pós graduação em Banco de Dados e Business Intelligence (2016). Atualmente é assessor de unidade estratégica - Banco do Brasil - Direção Geral. Tem experiência na área de Ciência da Computação, com ênfase em Arquitetura de Sistemas de Com-



MARCELO LADEIRA Possui graduação em Engenharia Mecânica pela Universidade de Brasília(1976), especialização em Curso Avançado de Operação de Sistema Hidrotérmico pela Universidade Federal do Rio de Janeiro(1981), mestrado em Análise de Sistemas e Aplicações pelo Instituto Nacional de Pesquisas Espaciais(1981), doutorado em Computação pela Universidade Federal do Rio Grande do Sul(2000) e pós-doutorado pela Universidade de Lisboa(2004). Atualmente é Pro-

fessor Adjunto 2 da Universidade de Brasília, Membro de Comitê Consultivo do Ministério da Ciência, Tecnologia, Inovações e Comunicações, Revisor de periódico da Journal of Information and Data Management - JIDM e Membro de comitê assessor do Conselho Nacional de Desenvolvimento Científico e Tecnológico. Tem experiência na área de Ciência da Computação, com ênfase em Metodologia e Técnicas da Computação. Atuando principalmente nos seguintes temas:Inteligência Artificial, raciocínio probabilístico, diagramas de influências, redes bayesianas, Teoria da decisão e Representação de conhecimento incerto.