

# Generating maximum prime patterns using Benders decomposition and Apriori algorithm

Hany Osman

Systems Engineering Department  
King Fahd University of Petroleum & Minerals, KFUPM  
Dhahran, 31261, Kingdom of Saudi Arabia.  
[hanyosman@kfupm.edu.sa](mailto:hanyosman@kfupm.edu.sa)

## Abstract

Incorporating data mining tasks in different levels of planning has become an essential tactic in business, industry, and other sectors. The rationales for implementing a data mining task, such as classification, significantly increase if the techniques used in classification provide optimal results. Logical Analysis of Data (LAD) is a classification approach known for its promising accuracy in classification and its capabilities in providing interpretable patterns. The main challenge in implementing LAD is the pattern generation problem. In this study, the pattern generation problem is solved to optimality to find maximum prime patterns. The proposed approach incorporates Benders decomposition and Apriori algorithm to generate prime patterns with high coverage from past observations. These patterns are then employed to build LAD classifiers that are used to assign class labels to unseen observations. Computational experiments conducted on seven public datasets show that results of LAD classifiers, established by using the proposed pattern generation algorithm, surpassed results of six machine learning algorithms implemented in IBM SPSS Modeler.

**Keywords:** Data mining; Logical analysis of data; Pattern generation; Benders decomposition; Apriori algorithm;

## 1. Introduction

Machine learning algorithms and data mining have received a significant attention in different sectors during the last two decades. This is because of the great achievement accomplished by their integration in decision making that has led to reaching better outcomes. Logical analysis of data (LAD) is a machine learning algorithm used for classification. The algorithm aims at finding the patterns existing in two-class datasets. LAD has been implemented in different applications belonging to engineering, medical, and business fields, Lauer et al. (2002), Alexe et al. (2003), Alexe et al. (2005), Bruner et al. (2007), Reddy et al. (2008), Kim et al. (2008), Hammer et al. (2010), Kogan et al. (2010), Hammer et al. (2012), Mortada et al. (2011a), Mortada et al. (2011b), Mortada et al. (2014), Ragab et al. (2016a), Ragab et al. (2016b), Ragab et al. (2017), and Bruni et al. (2018). LAD has been proposed in Crama et al. (1988) based on theory of Boolean functions. LAD consists of three main steps. The first step is called binarization. This step is required to transform nonbinary features to binary features. The second step, which is the most critical one, is called pattern generation. In this step, the patterns required to distinguish between positive and negative classes are generated. Generating patterns with high accuracy of classification is very critical to LAD performance. The difficulties associated with solving the pattern generation problem come from the combinatorial nature of this problem. The third step is called theory formation. In this step, the LAD classifier is established by using the generated patterns. The full implementation of LAD is presented in Boros et al. (2000).

The paper is organized as follows. A brief literature review on approaches developed to solve the pattern generation problem is given in section 2. Since the proposed decomposition approach is applied to Bonates et al. (2008) model, the model is discussed in section 3. The proposed decomposition algorithm for generating maximum prime  $\alpha$ -pattern

is presented in section 4. A comparative study between LAD classifiers, obtained by the proposed algorithm, and other machine learning algorithms is presented in section 5. Summary and conclusions of the paper are given in section 6.

## 2. Literature review on the pattern generation problem

The literature of LAD presents different approaches for solving the pattern generation problem. The problem can be represented by mathematical programming models and solved by commercial solvers, Bonates (2007), Bonates et al. (2008), Ryoo et al. (2009), Hansen et al. (2011), Guo et al. (2012), Caster et al. (2016), and Chou et al. (2017). Heuristic techniques for solving the pattern generation problem is developed in Boros et al. (2000), and Alexe et al. (2006). A comprehensive review on theory and applications of LAD is presented in Lejeune et al. (2018).

Pattern generation by using simple heuristics are proposed in Boros et al. (2000). Two heuristics are presented; a top-down and a bottom-up approaches. A more advanced heuristic is proposed in Alexe and Hammer (2006) to generate all spanned patterns in a dataset in a polynomial time. The algorithm employs Black and Quine consensus method in finding the prime implicants of Boolean function. Mixed binary integer programming models are proposed in Ryoo and Jang (2009). The purpose of the model is to generate strong prime patterns and strong spanned patterns. More compact binary models are proposed in Guo and Ryoo (2012). These models have the advantage of a fewer number of binary variables as compared to the models presented in Ryoo and Jang (2009).

An approach for generating a pattern covering a particular observation is introduced in Bonates (2007), Bonates et al. (2008). This kind of patterns is called maximum  $\alpha$ -pattern, where  $\alpha$  is a particular positive or negative observation. The generated pattern is maximum because it has maximum coverage from the  $\alpha$ -related class. Another kind of patterns that also known for its high coverage is the prime pattern, Bonates, et al. (2008). A pattern is called prime if it is minimal in terms of the number of literals, Boros et al. (2000). In addition, prime patterns are preferred to non-prime patterns because they are easier to be interpreted and employed in prognostic and diagnostic models.

In this paper, a decomposition algorithm is proposed to find maximum prime patterns. The algorithm is established based on combinatorial Benders decomposition technique introduced in Codato et al. (2014), and the Apriori property introduced in Agrawal et al. (1994). The proposed approach is applied to the mathematical model proposed in Bonates, et al. (2008) that generates maximum  $\alpha$ -pattern. The experiments conducted in this paper demonstrate that the proposed decomposition algorithm resulted in LAD classifiers possessing higher classification accuracy than six machine learning algorithms implemented in IBM SPSS Modeler 18.1.

## 3. Bonates et al., 2008 equivalent linear model

Bonates et al. (2008) introduced a nonlinear model and its equivalent linear model for generating maximum  $\alpha$ -pattern. A maximum positive (negative)  $\alpha$ -pattern is the positive (negative) pattern that covers a chosen positive (negative) observation  $\alpha$  and maximizes the coverage from the positive (negative) class. The proposed decomposition algorithm is applied on the equivalent linear model introduced in Bonates et al. (2008). The model is basically developed to linearize the nonlinear model. The equivalent model is presented here to provide a better understanding of the decomposition algorithm proposed for generating maximum prime  $\alpha$ -pattern.

Consider a binary data set  $\Omega$  that contains a set of positive observations  $\Omega^+$  and a set of negative observations  $\Omega^-$ . Also consider an observation  $\alpha$  that is chosen to generate the pattern. Observation  $\alpha$  is composed of binary features  $\alpha_j$ , where  $j= 1, 2, \dots, n$ . is an index referring to the binary features included in the data set  $\Omega$ . In the mathematical model given by equations (1) - (6), observation  $\beta$  is a positive observation that is composed of binary features  $\beta_j$ , while observation  $\gamma$  is a negative observation that is composed of binary features  $\gamma_j$ . The parameter  $w(\beta)$  is an integer parameter that is equal to  $|\{j: \beta_j \neq \alpha_j\}|$ . There are two variables in this formulation. The first one is  $y_j$  which is a binary variable that indicate whether the value of the  $j^{\text{th}}$  binary feature of the pattern is fixed to the value of  $\alpha_j$ . The second variable is  $z_\beta$  which is an auxiliary variable used to replace the nonlinear term  $\prod_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n \bar{y}_j$ , where  $\bar{y}_j = 1 - y_j$ . From this nonlinear term we see that in order to cover observation  $\beta$  by the pattern being generated, all  $y_j$  for which  $\beta_j \neq \alpha_j$

should be equal to 0. So, if  $\prod_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n \bar{y}_j = 1$ , it means that observation  $\beta$  is covered by the pattern. The pattern will be a maximum pattern if  $\sum_{\beta \in \Omega^+ \setminus \{\alpha\}} \prod_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n \bar{y}_j$  is maximized. This is represented by the objective function given in equation

(1). Constraint (2) ensures that the generated pattern does not cover any observation from the negative class. Constraints (3) and (4) are the auxiliary constraints required to achieve equivalence between the auxiliary variable  $z_\beta$  and the nonlinear term  $\prod_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n \bar{y}_j$ . Constraints (5) and (6) represent the binary restrictions imposed on the variables.

$$\text{Max} \quad \sum_{\beta \in \Omega^+ \setminus \{\alpha\}} z_\beta \quad (1)$$

Subject to

$$\sum_{\substack{j=1 \\ \gamma_j \neq \alpha_j}}^n y_j \geq 1 \quad \text{for every } \gamma \in \Omega^- \quad (2)$$

$$w(\beta) z_\beta + \sum_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n y_j \leq w(\beta) \quad \text{for every } \beta \in \Omega^+ \setminus \{\alpha\} \quad (3)$$

$$z_\beta + \sum_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n y_k \geq 1 \quad \text{for every } \beta \in \Omega^+ \setminus \{\alpha\} \quad (4)$$

$$y_j \in \{0,1\} \quad \text{for every } j = 1, \dots, n, \quad (5)$$

$$z_\beta \in \{0,1\} \quad \text{for every } \beta \in \Omega^+ \setminus \{\alpha\} \quad (6)$$

The model given by equations (1) - (6) considers that  $\alpha$ -observation belongs to the positive class. A similar model can be formulated for negative  $\alpha$ -observations to generate negative maximum  $\alpha$ -patterns.

#### 4. Finding maximum prime $\alpha$ -pattern

The mathematical model given by equations (1) - (6) do not guarantee generating prime patterns. Prime patterns are known for their generality and high coverage, Bonates et al. (2007). Therefore, they can be used to build accurate classifiers. In order to generate a prime pattern, the number of terms included in the pattern should be minimized. Consequently, one more objective is considered in the algorithm proposed to generate maximum prime patterns. This objective function aims at minimizing  $\sum_{j=1}^n y_j$ . In this section, a heuristic algorithm for generating maximum prime  $\alpha$ -patterns is discussed. The algorithm is presented for the case of generating maximum prime positive patterns. Maximum prime negative patterns can be generated similarly.

The algorithm is established based on combinatorial Benders decomposition proposed in Codato et al., (2004), and the Apriori property proposed in Agrawal et al., (1994). The algorithm is iterative and composed of a master problem and a subproblem. A lower bound representing the coverage of the pattern is recorded and updated at each iteration only if a better bound is found.

Equations (7) - (10) shows the Benders master problem that finds the minimum number of attributes required to form the prime positive  $\alpha$ -pattern. After solving the master problem, the Benders subproblem given in equations (11) - (15) is solved to maximize the coverage of the generated  $\alpha$ -pattern from the positive observations. The lower bound is updated if the coverage of the new pattern is higher than the recorded lower bound. The subproblem favors that all  $z_\beta$  to be equal to 1 which means covering all positive observations. Values of  $y_j$ , that are equal to 1, found by solving the master problem make all  $z_\beta$  equals 1 except any positive observation  $\beta$  that has at least one feature  $j$  for which  $\beta_j \neq \alpha_j$  (indicating to a difference between observation  $\beta$  and the generated pattern). Consequently, there is no need to solve

the subproblem by a commercial solver. Instead, a simple check can be implemented. For a positive observation  $\beta$ , if for those  $y_j = 1$  their associated  $\beta_j \neq \alpha_j$  then this observation  $\beta$  is not covered by the pattern and is covered otherwise.

At each iteration  $q$ , the combinatorial cut given by equation (9) is added to the master problem. This will force the master problem to generate a different 0-1 combination of  $y_k$ , which means generating a new  $\alpha$ -pattern. The subproblem is checked again to check whether the new pattern has better coverage. The lower bound is updated if better coverage is found. The algorithm stops when the master problem has no feasible solution. The resulted solution is optimal because the entire search space is explored.

Benders master problem

$$\text{Min} \quad \sum_{j=1}^n y_j \quad (7)$$

Subject to

$$\sum_{j=1}^n y_j \geq 1 \quad \text{for every } \gamma \in \Omega^- \quad (8)$$

$$\sum_{y_j^q=0}^n y_j - \sum_{y_j^q=1}^n (1 - y_j) \geq 1 \quad \forall q \in Q \quad (9)$$

$$y_j \in \{0,1\} \quad \text{for every } j = 1, \dots, n, \quad (10)$$

Benders subproblem

$$\text{Max} \quad \sum_{\beta \in \Omega^+ \setminus \{\alpha\}} z_\beta \quad (11)$$

Subject to

$$w(\beta) z_\beta + \sum_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n y_j \leq w(\beta) \quad \text{for every } \beta \in \Omega^+ \setminus \{\alpha\} \quad (12)$$

$$z_\beta + \sum_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^n y_j \geq 1 \quad \text{for every } \beta \in \Omega^+ \setminus \{\alpha\} \quad (13)$$

$$y_j = y_j^* \quad \text{for every } j = 1, \dots, n, \quad (14)$$

$$z_\beta \in \{0,1\} \quad \text{for every } \beta \in \Omega^+ \setminus \{\alpha\} \quad (15)$$

The Apriori property is a common property in data mining for finding frequent itemsets in market basket analysis. The Apriori property states that if an itemset is found to have a frequency  $f$ , then adding another item to this itemset will not make it more frequent. The same property can be applied to patterns and their coverage. The master problem in its classical combinatorial constraint (9) does not consider that. The combinatorial cut can be updated by employing the Apriori property in its formulation. Given a pattern  $p$  that is composed of few features  $j$  and has a coverage  $Z$ , adding one or more feature to this pattern will not increase its coverage. For example, consider a dataset containing five features, the master problem generates a prime pattern composed of the two features  $j=1$  and  $j=2$ . This means that  $y_1=y_2=1$  and the rest  $y$ 's are equal to zero. Applying the classical combinatorial cut may result in a new pattern that just adds any of  $y_3$  or  $y_4$  or  $y_5$  to the generated pattern. From the Apriori property, it is confirmed that the resulted pattern will not have coverage (lower bound) better than the pattern composed of only the first and second features,  $j=1$  and  $j=2$ . So, the master problem should not generate any new pattern that includes previously resulted combinations of  $y_j=1$ , i.e.  $y_1=y_2=1$  in this example. The Apriori property can be represented in the master problem by

omitting the first term in the left-hand side of the classical combinatorial cut (15). Equation (21) shows the updated combinatorial cut that makes use of the Apriori property. The updated combinatorial cut reduces the feasible area investigated in the master problem and results in faster optimal solutions.

$$\sum_{y_k^q=1}^K (1 - y_k) \geq 1 \quad \forall q \in Q \quad (21)$$

The sets of positive and negative patterns generated by the proposed approach are then used to build the classifier shown in equation (22). This classifier is used to classify unseen observations to be either belong to the positive class ( $\Delta$  is positive) or to the negative class ( $\Delta$  is negative).  $P_k$  is a parameter that is equal to 1 if the unseen observation is covered by positive pattern  $k$  and 0 otherwise, while  $N_l$  is a parameter that is equal to -1 if the unseen observation is covered by negative pattern  $l$  and 0 otherwise. The weights  $w_p^+$  and  $w_n^-$  given to patterns in this equation represent their prevalence, Ryoo et al. (2009).

$$\Delta = \sum_{k=1}^K w_k^+ P_k + \sum_{l=1}^L w_l^- N_l \quad (22)$$

## 5. Comparison to other machine learning algorithms.

In this section, the classifiers resulted from the pattern generated by the proposed approach are compared to six machine learning algorithms implemented in IBM SPSS Modeler 18.1. The cross-validation  $k$ -fold algorithm with  $k$  equals 10 is used to validate each classification model resulted by these algorithms. IBM SPSS Modeler 18.1 is used to generate the 10 folds, where same folds are used in testing and training phases of the proposed algorithm and in the six data mining algorithms. The CPU used in these experiments is Intel(R) Xeon(R), 3.5GHz with 32G RAM.

The data sets used in this comparative study are described in table 1. Observations with missing values are dropped from the datasets. The fourth column of the table shows the number of binary features after implementing the binarization step. The algorithm used for binarization is the one presented in Mortada et al. (2011a). The last column of the table shows the size of the support set used in the pattern generation problem. Boros et al. (2000) set covering problem is used to find the support set.

Table 1: The seven public datasets used in the comparative experiments

Dataset*	Number of observations	Number of features	Number of binary features	Size of the support set
Housing	505	13	1245	16
SPECTF	267	44	992	10
Ionosphere	351	34	2308	11
Congressional Voting	233	16	16	8
Breast Cancer	683	9	81	13
Pima	786	8	857	21
Liver Disorders	345	6	269	17

\* <https://archive.ics.uci.edu/ml/datasets.html>

Table 2 presents results of the maximum prime  $\alpha$ -patterns generated by the proposed decomposition algorithm. As shown in the third column of the table, the time elapsed in generating prime  $\alpha$ -patterns is ranging from few seconds to less than 1.2 minute. The last column shows the average size of the pattern in terms of the number of literals, while the second to last column shows the average coverage of the patterns from the training set.

Table 2: Results of the proposed approach

Dataset	Accuracy	Time (min)	#of Patterns	Coverage	Size
Housing	89.5±2.81	1.3±0.09	78±2.41	32.2±1.84	4.61±0.11
SPECTF	79.6±7.28	1.1±0.02	62.5±3.54	15.93±1.09	4.74±0.08

Ionosphere	92.44±2.63	0.53±0.08	38.2±2.15	29.17±2.03	3.97±0.1
Voting	94.31±5.32	0.10±0.05	13.5±2.33	37.2±4.8	3±0.22
Breast cancer	95.41±1.94	0.24±0.09	42.6±2.07	72.98±4.72	3.55±0.15
Pima	71.81±3.94	1.2±0.090	190.9±4.1	523.37±0.65	6.93±4.07
Liver Disorders	71.17±9.52	0.6±0.04	95.1±4.34	8.54±0.44	4.98±0.07

\* Results are given in the format; mean ± standard deviation.

The seven public data sets described in table 1 are used to compare LAD classifier results, obtained by using the proposed decomposition algorithm, to six machine learning algorithms. These algorithms are Support Vector Machine (SVM), Neural Networks (NN), Bayes Networks, Decision Trees C.5, K-Nearest Neighbours (KNN), and Random Forests (RF). Parameters of these algorithms are set at their default values in IBM SPSS Modeler 18.1. For these seven datasets, table 3 shows the accuracy of classification resulted by using LAD and IBM Modeler's implementation of these six machine learning algorithms. The accuracy is calculated as the percentage of correctly classified observation from the total number of observations. As seen in the table, LAD outperforms each of these six algorithms in classifying the testing datasets, where it provides the highest accuracy in four datasets; Housing, Ionosphere, Voting, and Liver Disorders. Also, LAD provides close to the highest accuracy in the datasets; SPECTF, Breast cancer, and Pima. In addition to this high performance, LAD provides interpretable patterns which is a key advantage of LAD. The performance of SVM, NN, and RF is very competitive but each of these algorithms is missing the advantage of providing interpretable patterns or decision rules. Interpretable patterns are highly required in real life problems especially in prognosis applications when it is required to guide a process or a system toward a specific class. For example, guiding a CNC machine to keep producing good parts by controlling the speed and feed rate, and other input features.

Table 3: Comparison of the proposed approach to other classifiers

Dataset	Support vector machines	Neural networks	Bayes Net	Decision trees C5	K-nearest neighbour	Random Forest	Proposed LAD approach
Housing	88.44±5.3	86.37±5.56	79.47±3.86	82.36±6.51	85.89±4.67	87.45±4.68	<b>89.5±2.81</b>
SPECTF	<b>82.45±6.11</b>	81.99±6.85	63.53±14.7	81.69±8.88	80.12±4.73	81.99±7.82	79.6±7.28
Ionosphere	87.33±6.23	91.04±5.35	59.74±7.85	89.67±3.7	82.74±9.39	92.45±4.49	<b>92.83±2.63</b>
Voting	87.27±3.19	87.93±3.74	85.5±3.92	86.78±2.2	86.82±3.78	87.06±4.59	<b>94.31±5.32</b>
Breast Cancer	96.94±1.76	96.41±2.28	95.39±3.9	95.31±2.91	<b>97.87±1.54</b>	96.87±1.65	95.41±1.94
Pima	76.05±3.52	75.98±3.18	<b>76.38±4.69</b>	73.73±3.41	73.01±4.12	73.58±5.06	71.81±3.94
Liver Disorders	64.22±5.86	65.26±5.75	54.84±9.28	63.92±4.87	60.44±7.8	67.85±10.43	<b>71.17±9.52</b>

## 6. Summary and Conclusions

Logical analysis of data is a machine learning algorithm that has shown competitive results when applied to binary classification problems. Among the steps required to implement LAD, solving the pattern generation problem is the most challenging one. In this paper, a new decomposition approach is developed by employing the Apriori property in the combinatorial Benders decomposition approach to generate optimal maximum prime  $\alpha$ -patterns. Prime patterns are known for their high accuracy of classification. Computational experiments are conducted to compare results of the model proposed to results of six machine learning algorithms implemented by IBM SPSS Modeler 18.1. Seven public datasets are used in these computational experiments. The experiments demonstrated a promising performance of the proposed approach.

## Acknowledgment

The author would like to acknowledge the support provided by the Deanship of Scientific Research (DSR) at King Fahd University of Petroleum & Minerals (KFUPM) for funding this work through project No. SR141010.

## References

- Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. In *Proceedings of the 20th International Conference on Very Large Data Bases, VLDB, Santiago, Chile*, Santiago, Chile.
- Alexe, G., Alexe, S., Axelrod, D.E., Hammer, P.L., Weissmann, D. 2005. Logical analysis of diffuse large B-cell lymphomas. *Artif. Intell. Med.* 34, 235–267.
- Alexe, S., Blackstone, E., Hammer, P.L., Ishwaran, H., Lauer, M.S., Snader, C.E.P., 2003. Coronary risk prediction by logical analysis of data. *Annals of Operations Research*. 119, 15–42.
- Alexe, G., & Hammer, P. L. (2006). Spanned patterns for the logical analysis of data. *Discrete Applied Mathematics*, 154(7), 1039–1049. <http://doi.org/10.1016/j.dam.2005.03.031>
- Bonates, T. O. 2007. Optimization in logical analysis of data. PhD thesis, Rutgers. The State University of New Jersey.
- Bonates, T.O., Hammer, P.L., Kogan, A. 2008. Maximum patterns in datasets. *Discrete Appl. Math.* 156, 846–861.
- Boros, E., Hammer, P.L., Ibaraki, T., Kogan, A., Mayoraz, E., Muchnik, I., 2000. An implementation of logical analysis of data. *IEEE transactions on knowledge and data engineering* 12, 292–306
- Brauner, M.W., Brauner, N., Hammer, P.L., Lozina, I., Valeyre, D., 2007. Logical analysis of computed tomography data to differentiate entities of idiopathic interstitial pneumonias. *Data Mining in Biomedicine*. Springer, pp. 193–208.
- Bruni, R., Bianchi, G., Dolente, C., Leporelli, C., 2018. Logical analysis of data as a tool for the analysis of probabilistic discrete choice behavior. *Computers and Operations Research*. 000, 1-11
- Caserta, M., Reiners, T., A pool-based pattern generation algorithm for logical analysis of data with automatic fine-tuning. *European Journal of Operational Research* 248 (2016) 593–606.
- Chou, C. A., Bonates, T.O., Lee, C., Chaovalitwongse, W.A., 2017. Multi-pattern generation framework for logical analysis of data. *Annals of Operations Research*. 249(1), 329–349
- Codato G., Fischetti M. (2004) Combinatorial Benders' Cuts. In: Bienstock D., Nemhauser G. (eds) *Integer Programming and Combinatorial Optimization*. IPCO 2004. Lecture Notes in Computer Science, vol 3064. Springer, Berlin, Heidelberg
- Crama, Y., Hammer P.L., Ibaraki, T. (1988) Cause-effect relationship and partially defined boolean functions. *Annals of operations research*. 16, 299–325.
- Guo, C., Ryoo, H.S., 2012. Compact MILP models for optimal and Pareto-optimal LAD patterns. *Discrete Applied Mathematics*. 160, 2339–2348.
- Hammer, P. L., Kogan, A., Lejeune, M. A. 2010. Reverse-engineering country risk ratings: a combinatorial non-recursive model. *Annals of Operations Research*, 188(1), 185–213
- Hammer, P.L., Kogan, A., Lejeune, M.A., 2012. A logical analysis of banks' financial strength ratings. *Expert Systems with Applications*. 39, 7808–7821.
- Hansen, P., Meyer, C., 2011. A new column generation algorithm for logical analysis of data. *Annals of Operations Research*. 188, 215–249.
- Kim, K., Ryoo, H.S., 2008. A LAD-based method for selecting short oligo probes for genotyping applications. *OR Spectrum* 30, 249–268.
- Kogan, A., Lejeune, M. A. 2010. Combinatorial methods for constructing credit risk ratings. *Handbook of Quantitative Finance and Risk Management*. 639-664. Springer
- Lauer, M. S., Alexe, S., Snader, C. E. P., Blackstone, E. H., Ishwaran, H., Hammer, P. L. 2002. Use of the logical analysis of data method for assessing long-term mortality risk after exercise electrocardiography. *Circulation*, 106, 590–685.
- Lejeune, M., Lozin, V., Lozina, I., Ragab, A., Yacout, S., 2018 Recent advances in the theory and practice of Logical Analysis of Data. *European Journal of Operational Research*. <https://doi.org/10.1016/j.ejor.2018.06.011>
- Mortada, M.-A., Yacout, S., Lakis, A., 2011. Diagnosis of rotor bearings using logical analysis of data. *Journal of Quality Maintenance in Engineering*. 17, 371–397.
- Mortada, M.-A., Yacout, S., Lakis, A., 2011. Diagnosis of rotor bearings using logical analysis of data. *Journal of Quality Maintenance in Engineering*. 17, 371–397.
- Mortada, M.-A., Yacout, S., Lakis, A., 2014. Fault diagnosis in power transformers using multi-class logical analysis of data. *Journal of Intelligent Manufacturing* 25, 1429–1439.
- Ragab, A., Ouali, M.-S., Yacout, S., Osman, H. 2016. Prognostics of Multiple Failure Modes in Rotating Machinery Using a Pattern-Based Classifier and Cumulative Incidence Functions. *Journal of Intelligent Manufacturing*. DOI: 10.1007/s10845-016-1244-8.
- Ragab, A., Ouali, M.-S., Yacout, S., Osman, H. 2016. Remaining useful life prediction using prognostic methodology based on logical analysis of data and Kaplan–Meier estimation. *Journal of Intelligent Manufacturing*. 27(5), 943–958

- Ragab, A., Ouali, M.-S., Yacout, S., Osman, H. 2017. Pattern-based prognostic methodology for condition-based maintenance using selected and weighted survival curves. *Quality and Reliability Engineering International*. DOI: 10.1002/qre.2127
- Reddy, A., Wang, H., Yu, H., Bonates, T.O., Gulabani, V., Azok, J. 2008. Logical analysis of data (lad) model for the early diagnosis of acute ischemic stroke. *BMC Medical Informatics and Decision Making*, 8(30).
- Ryoo, H.S., Jang, I.-Y., 2009. MILP approach to pattern generation in logical analysis of data. *Discrete Applied Mathematics* 157, 749–761.

### **Biography**

Hany Osman received the Ph.D. degree in industrial engineering from Concordia University, Montréal, QC, Canada, in 2011. He is an Assistant Professor with the College of Computer Sciences and Engineering, King Fahd University of Petroleum and Minerals, KSA. His research interests include data mining, operations research, and operations management. Specifically, his research focuses on investigating problems related to knowledge extraction from datasets, scheduling, inventory management, and transfer line balancing. He has developed efficient algorithms by using decomposition techniques and nature-inspired metaheuristics.