Proceedings of the International Conference on Industrial Engineering and Operations Management
Pilsen, Czech Republic, July 23-26, 2019

Potential for PRIM based classification: A literature review

Rym Nassih
AMIPS Research team - EMI
University Mohamed V in Rabat
Morocco
rymnassih@research.emi.ac.ma

Abdelaziz Berrado
AMIPS Research team - EMI
University Mohamed V in Rabat
Morocco
berrado@emi.ac.ma

Abstract

This paper provides a critical review of Patient Rule Induction Method and other bump hunting algorithms and their applications. We also give an overview about interpretability in several key supervised data mining algorithms. This allows for exploring the potential for using PRIM, with its interpretation capability, as a core technology towards building a highly accurate and interpretable classifier in a mixed data space.

Keywords
Supervised Learning, PRIM, Bump Hunting, Interpretability

1. Introduction

With the advent of supervised learning research, several data mining algorithms have been developed. Most of the available classification algorithms offer a high accuracy for most problems but are unable to provide insight into predictive structure of supervised learning algorithms. Furthermore, interpretability is necessary for decision making, especially in the context of sensitive application areas such as healthcare, security or sociology. Black box models still need to be demystified.

Machine learning models have reached, today, a high level of accuracy. They have demonstrated a great success in learning complex patterns that enable them to make predictions. Thus, using such models for predictions requires a high amount of attention to achieve a satisfied level of interpretability. However the notion of interpretability leads to confusion since it is unclear how the wide array of proposed interpretation methods are related and what common concepts can be used to evaluate them.

Assessing model's trust was discussed by many authors. In [1], interpretability is considered as a way to the model's trust by providing a transparent reasoning on how it works. Also in [2], the same idea was reported with insisting on the importance of intelligibility and modularity.

In addition to that, many authors addressed the interpretability through its ability to enable "action and reaction" by means of results exploratory. In the same context, in [3] and [4] authors proposed, respectively, Variables Importance plots and Partial Dependency plots as tools to provide insights into the nature of inputs-outputs relationships. Also in [5] [6] [7], authors used the model-agnostic approach. Authors in [39] propose a scoring grid for scoring interpretability measures of any machine learning (ML) system such as intelligibility, modularity and patterns discovery.

Authors in [8] claim that interpretability is not only a mathematical notion. They provide a comprehensive taxonomy of both the desiderata (trust, causality, transferability, informativeness and fair and ethical decision-making) and methods (in interpretability research. For this purpose, in [9], authors enumerated several terms associated with interpretability (understandability, comprehensibility, justifiability and usability).

In our case we focus on the interpretability using supervised algorithms, by considering three based methods: the tree based method, the rule based method and the bump hunting method.

In this paper, we attempt to highlight a bump hunting algorithm called Patient Rule Induction Method. To do so, we begin with a review of the interpretability challenges in supervised learning problems. Next we provide a state of the art of key supervised learning methods including tree based methods, rule based methods and bump hunting methods such as PRIM. In section 4, we carry out a critical review about PRIM, bump hunting algorithms and their applications which lead us to guidelines for future work aiming to integrate an adapted version of PRIM for classification.

2. State of art of key ensemble supervised methods

In order to explore the potential for interpretability in key supervised learning algorithms we briefly present, in this section, Tree Based methods such as IDE and its improvements, CART, Random Forest and Gradient Boosting for classification. We also present rule based methods including Rulefit for classification, Associative Classification and RCAR (Regularized Classification with Association Rules). Finally we provide an overview on Bump hunting with Data Surveyor and PRIM.

2.1 Tree based method

2.1.1 Iterative Dichotomizer and its improvements

ID3 or Iterative Dichotomizer, was the first of three Decision Tree implementations developed by Ross Quinlan [10]. It builds a decision tree for the given data in a top-down fashion, starting from a set of objects and a specification of properties. ID3's main disadvantages are the fact that data may be over-fitted or over-classified, if a small sample is tested. In addition to that, only one attribute at a time is tested for making a decision and it does not handle numeric attributes and missing values.

C4.5 is an improved version of ID3 [11], in addition of what ID3 is, it also accepts both continuous and discrete features, it handles incomplete data points and it solves over-fitting problem by using bottom-up pasting procedure commonly known as "pruning". But unfortunately, C4.5 can construct empty branches with zero values and over-fitting can happen if the data is noisy.

For those reasons, Quinlan [11] tried to improve C4.5 and came out with C5.0, the most recent iteration which stand out with the speed comparing to C4.5 and the fact that it gets similar results to C4.5 with considerably smaller decision trees.

2.1.2 CART and Random Forest

Classification and Regression Trees (CART) [12] constructs binary trees. The splits are selected using the twoing criteria and the obtained tree is pruned by cost–complexity Pruning. CART can handle both numeric and categorical variables and it can easily handle outliers. The notable difference between CART and C4.5 is that CART constructs the tree based on a numerical splitting criterion recursively applied to the data, whereas C4.5 includes the intermediate step of constructing rule sets.

Developed by Breiman [13], Random Forest consists of a collection or ensemble of simple tree predictors, each capable of producing a response when presented with a set of predictor values. For classification problems, this response takes the form of a class membership, which associates, or classifies, a set of independent predictor values with one of the categories present in the dependent variable. This arbitrary number of simple trees, are used to determine the final outcome. For classification problems, the ensemble of simple trees vote for the most popular class.

Random Forest is generally a better model if the goal is for prediction, however for exploratory analysis, often when we just want to understand how the data is related in a tree hierarchy structure, a single tree (CART) is preferable

2.1.3 Gradient boosting for classification trees

It, first, started with AdaBoost algorithm which was introduced in [14] as the first successful boosting algorithm and was formulated as a gradient descent with a special loss function in it [15] [16].

Boosting consists in learning a collection of weak base learners, commonly classification trees. The procedure is to create modified version of the data and then to train the classification trees sequentially on it.

Gradient boosting attempts to construct an approximation of the function mapping the input variables to the response variable by minimizing a loss function over a training data set with no lack of values.

Random Forest are much easier to tune than GBM. There are typically two parameters in Random Forest: the number of trees and the number of features to be selected at each node. In addition to that Random Forest are harder to overfit than Gradient Boosting (if the data is noisy). In addition of that, training in Gradient Boosting generally takes longer because of the fact that trees are built sequentially.

2.2 Rule based method

2.2.1 Association classification

In associative classification, association rules are generated and analyzed for use in classification. It explores highly confident associations among multiple attributes and may overcome some constraints introduced by decision-tree induction, which considers only one attribute at a time

Authors in [18] and [19] give an exhaustive review on the association classification mining. Among them all, we present the typical associative classification methods. The first one, the Classification By Association (CBA) [20] consists of two parts, a rule generator (called CBA-RG), which is based on algorithm Apriori for finding association rules [21], and a classifier builder (called CBA-CB) which function is to organize the rules according to decreasing precedence based on confidence and then support.

And then Li at al in [22] came with the Classification based on Multiple Association Rules (CMAR) which outperforms CBA in terms of efficiency. It uses as a rule discovery the FP-growth approach and it ranks rules depending on the confidence, the support and the rule cardinality.

There are many other classification algorithm as MMSCBA [23] which adopts minimum support thresholds and thereby gives a better accuracy than the conventional associative classifier. Another one would be the Adaptative-Support Association Rule Mining (ASARM)[24] for recommender systems.

Having, as a result, a huge set of mined rules can lead to confusion in the interpretability of the models. Thus, Berrado et al. in [25] proposed to use metarules to organize and group the redundant rules by revealing the relationships between them.

2.2.2 Rulefit

A rule-based ensemble method named Rulefit was introduced by Friedman [17], and have competed with the previous tree based algorithms. RuleFit is a so-called ensemble method, it combines the predictions of a large number of simple models. Those simple models are referred to as base learners or weak learners, as their predictions often perform only marginally better than random guessing. By combining predictions of a large number of base learners, ensemble methods perform better than any of their constituent members. The base learners in a RuleFit model are prediction rules: statements of the form: if [condition], then [prediction]. The condition specifies a set of values of predictor variables, and the prediction specifies the expected increase or decrease in the criterion variable, when an observation satisfies the specified condition. Rules in a RuleFit model are derived from classification and regression trees.

Rulefit procedure is constructed as a regularized linear combination of the original features and decision rules which are extracted from decision trees. This combination lead to a good improvement of accuracy and interpretability.

2.2.3 Regularized Class Association Rules Algorithm or Classification (RCAR)

In [26], Azmi et al. introduced a new classification algorithm based on Regularized Class Association Rules (RCAR). RCAR occurs in three steps. The first one is mining an exhaustive set of CARs according to predetermined value of minimum support and confidence. Then, we fit the Lasso regularized logistic regression model on the rule space. And the third step consists in using metarules to analyze the retained rules in the model.

2.3 Bump hunting based methods

Friedman and Fisher [27] introduced Bump Hunting, which aims at finding regions in the input space with relatively high (low) target value. It facilitates for the user to optimize the value of the target value by choosing or selecting the input variable of his choice. These regions have a rectangular shape and are described by simple rules of the type if: condition-1 and….and condition-n then: estimated target value. The box construction is a rule induction where the goal is to produce a box B within which the target mean is as large as possible. We discuss two approaches to this problem.

2.3.1 Patient Rule Induction Method (PRIM)

Patient Rule Induction Method (PRIM) is a bump hunting algorithm [27] which consists in two phases. The first one, is the top down peeling in which we construct the boxes and the second one is the bottom up pasting in which we optimize the result. We give an overview about PRIM in the next section. However, the most important advantage of

PRIM among others is its patience. Indeed, in each step only a small part of the data points in the current box is peeled of, hence the term patient rule induction. Thereby, if we compare PRIM to CART, CART fragments the data much faster than PRIM, since CART makes binary trees on average half of the data is removed at each step in the search. So we can run easily out of data.

2.3.2 Data Surveyor

Data surveyor in another bump hunting algorithm introduced by Holsheimer.at al [28] which box construction differs from PRIM since in PRIM we begin with one box and in data surveyor we select ω boxes wherein we select the best sub-boxes. Like in PRIM each eligible sub-box is defined on a single variable, but in a different manner. The way the sub-boxes are constructed is potentially greedier than in PRIM, since it may leads to a more rapid fragmentation of the data. Therefore, rather than peeling of a small part of the numeric variables, the algorithm goes directly for intervals with the highest target mean.

The important difference with PRIM regarding the search strategy is that PRIM uses a hill-climber, which means that it only considers the best peeling action on the current box. Whereas the Data Surveying search algorithm employs a beam search, in other words at each level in the search the best ω sub-boxes are considered.

3. Critical review of patient rule induction method

3.1 An overview of PRIM

The Patient Rule Induction Method suggested by Friedman and Fisher [27] is referred to as a bump hunting algorithm. It is used to find regions (one by one) in the input variable space that are associated with the highest or lowest mean value for the outcome. These rules have the following form: if condition1 and…and conditionK, then estimated mean outcome value.

These conditions can use numeric or categorical attributes. For continuous attributes, a condition will have the following form:

variable < value, or value1 < variable < value2

For categorical attributes, conditions have the following form:

variable = value variable = value1 or…or valuem

These rules are hypercube in the input variable space and take the form of rectangles in a two-dimensional space, hence the name box.

PRIM algorithm consists of the iteration of two phases: a top down peeling and a bottom up pasting. The top down peeling is the phase in which we construct the box. PRIM starts with a box containing all the given observations. At each step we remove a sub-box (a small portion of observations) that contains the largest mean target value. This candidate sub-box is defined depending on the type of the variable (numeric or categorical). The portion of data to be removed at each peel can be controlled by the data analyst and is specified in the tuning parameters as α, usually it's 5%.

The final box found after peeling may not be optimal because of past greedy suboptimal choices. To recover from these mistakes, PRIM uses the second phase: the bottom up pasting. It consist of expanding the box by iteratively enlarge its boundaries as long as the outcome's mean increases. The boxes found at end are a list of rules.

To avoid overfitting, PRIM uses cross-validation. The data set is randomly divided in a learning set and a training set, typically the learning set is taken to be twice the size of the test set. Then, the procedure is applied on the learning set with a very small value of the initial support β0, allowing for very small box supports. After this, we use the training set to estimate the output mean in each successive box induced on the training set. If we notice a significant difference in this output mean, then we have an overfitting and it is recommended not to trust such boxes.

Number of tools are also provided by PRIM to post-process or inspect the rules discovered such as the removal of redundant variables, the determination of inter-box dissimilarity and plotting relative frequency ratio plots. [27]

3.2 Related Work

PRIM has been formalized by Polonik & Wang in [29], they have proposed to replace the bottom up pasting with jittering. So rather than just adding small sets we simultaneously add and substract a box from the candidate boxes as long as we can increase the average of the box. Even if the complexity of the algorithm is somewhat increased by doing so, it has the advantage to enable us to find a characterization of the boxes and thus we can seek large sample results for the PRIM outcomes.

In [30], Chong et al. introduced a flexible patient rule induction method which was specially developed to deal with ordinal discrete variables. The authors applied the flexible PRIM and concluded that it is a good alternative method to process optimization when process variables under interest are in discrete type and much noisy information is contained in the data.

Authors in [31], proposed a Bayesian-assisted PRIM algorithm that covers regions in the feature space based on Bayes factor values and marginal posterior probabilities. This Bayesian model-assisted PRIM facilitates the implementation of PRIM and avoids the frequent user interactions, thus providing more accurate approximations to regions where the response variable has the maxima.

In [33], authors mixed PRIM algorithm with Local Sparse Bump Hunting algorithm [34] to develop an extension called fastPRIM under normality conditions and then combine it with the Principal Component (PC) rotation of the predictor space alone. This technique showed, at the end of the paper, that as long as the principal components are not being selected in priorto modeling the response, then these improved variables can produce more accurate mode characterizations.

Among the competitors to PRIM, CART is the one using the least greedy strategy. Therefore, many works have been done to compare the performance of the two. In [27], authors shows that PRIM yields better results, since CART makes binary tree on average, half of the data is removed at each step in the search, thereby it fragments the data much faster than PRIM. Whereas authors in [32], compared the performance of the two algorithms on a medical database and discovered that PRIM failed at discovering a subgroup. One cause is that PRIM doesn't handle in an optimal way the ordinal variables.

Other applications of PRIM has been done in the literature, and it all leads us to guidelines for future work.

3.3 Review of PRIM's application

In order to stress PRIM's shortcomings and limitations in high dimensional data, we briefly present some applications in the literature.

To see how PRIM performs comparing to logistic regression and CART, Abu-Hanna et al. in [35] applied, first, PRIM and logistic regression to select high-risk subgroups in very ederly dataset patients. They obtained satisfied results for PRIM as it arrives at "bumps" at different regions of the feature space than those found by the logistic regression model. And in their second work [32], the authors applied PRIM and CART on the same database and noticed that CART performed best because in one peel PRIM did the "wrong" choice: since it can peels just a small amount of data (α) if PRIM's choice is not the optimal one it may not be able to recover from its mistake. However the authors aim was to reach the highest level of accuracy, without given attention to the interpretability side of the model.

Also in [36] [37] [38], authors applied PRIM on high dimensional data. The common point between all the studies, is the intensive user interaction required by PRIM even if it gives a large amount of tools for diagnostic to the data analyst. Nevertheless, from those paper we notice that PRIM is weak in providing us with a satisfied degree of interpretability.

4. Conclusion and Guidelines

This paper highlights the intense need of interpretability in today's ML. Indeed, ML have reached a high level of accuracy but still miss the same level of interpretability. Therefore, the numerous data mining algorithms could give different models with a various complexity.

After presenting key classification supervised data mining algorithms, we have focused on PRIM, a bump hunting algorithm. Although PRIM presents several advantages, studies in literature have revealed several of its limitations. Researchers have highly participated in improving some PRIM's shortcomings [29] [30] [31] [33] [34]. Nevertheless, the huge amount of rules provided in high dimensional data still requires both the data analyst and the expert analysis and interaction. Thus, the interpretability relies on decision makers input and interaction.

This paper is a first step for us to set the problem of interpretability in PRIM and guides us to future work in the same perspective with the aim of integrating PRIM in a supervised learning context to give rise to interpretable classification models.

References

[1] Greg Ridgeway, David Madigan, Thomas Richardson, and John O'Kane. Interpretable Boosted Naïve Bayes Classification. KDD-98 Proceedings, page 4, in press

[2] Rich Caruana, Yin Lou, Johannes Gehrke, Paul Koch, Marc Sturm, and Noemie Elhadad. Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission. pages 1721–1730. ACM Press, 2015. in press

[3] Leo Breiman. Random Forests. Machine Learning, 45(1):5–32, October 2001. in press
[4] Jerome H. Friedman. Greedy Function Approximation: A Gradient Boostingn Machine. Annals of Statistics, 29:1189–1232, 2000. in press
[5] David Baehrens, Timon Schroeter, Stefan Harmeling, Motoaki Kawanabe, Katja Hansen, and Klaus-Robert Mu ller. How to Explain Individual Classification Decisions. Journal of Machine Learning Research 11 (2010) 1803-1831, page 29. in press
[6] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. pages 1135–1144. ACM Press, 2016. in press
[7] Sameer Singh, Marco Tulio Ribeiro, and Carlos Guestrin. Programs as Black-Box Explanations. arXiv:1611.07579 [cs, stat], November 2016. in press
[8] Zachary C. Lipton. The Mythos of Model Interpretability. arXiv:1606.03490 [cs,stat], June 2016. in press
[9] Adrien Bibal and Benoît Frénay. Interpretability of Machine Learning Models and Representations: An Introduction. Computational Intelligence, page 6, 2016. in press
[10] Quinlan, J.R. (1986). Induction of Decision Trees.*Machine Learning* 1:1, 81–106. in press
[11] Salzberg, Steven L. « C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers,
Inc., 1993 ». *Machine Learning* 16, nº 3 (septembre 1994): 235-40. in press
[12] L. T.G. Dietterich Ensemble methods in machine learning Multiple Classifier Systems, Springer (2000), pp. 1-15 in press
[13] L. Breiman Random forests Mach. Learn., 45 (1) (2001), pp. 5-32 in press
[14] Y. Freund, R.E. Schapire, et al.Experiments with a new boosting algorithm Proceedings of the ICML, 96 (1996), pp. 148-156 in press
[15] J.H. Friedman Stochastic gradient boosting Comput. Stat. Data Anal., 38 (4) (2002), pp. 367-378 in press
[16] J.H. Friedman Greedy function approximation: a gradient boosting machine Ann. Stat. (2001), pp. 1189-1232 in press
[17] J. Friedman, T. Hastie, R. Tibshirani Regularization paths for generalized linear models via coordinate descent J. Stat. Softw., 33 (1) (2010), p. 1 in press
[18] Sun, Yanmin, Andrew K C Wong, et IEEE Yang Wang. « An Overview of Associative Classifiers », s. d., 16. in press
[19] Thabtah, Fadi. « A Review of Associative Classification Mining ». The Knowledge Engineering Review 22, no 01 (mars 2007): 37 in press
[20] B. Liu, W. Hsu, Y. Ma, B. Ma Integrating classification and association rule mining Knowl. Discov. Data Min. (1998), pp. 80-86 in press
[21] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, A.I. Verkamo, *et al.* Fast discovery of association rules Adv. knowl. Discov. Data Min., 12 (1) (1996), pp. 307-328 in press
[22] W. Li, J. Han, J. PeiCmar: Accurate and efficient classification based on multiple class-association rules Proceedings of the IEEE International Conference on Data Mining, ICDM, IEEE (2001), pp. 369-376 in press
[23] L.-Y. Hu, Y.-H. Hu, C.-F. Tsai, J.-S. Wang, M.-W. Huang Building an Associative Classifier with Multiple Minimum Supports 5 (1) (2016), p. 528 in press
[24] W. Lin, S.A. Alvarez, C. Ruiz Efficient adaptive-support association rule mining for recommender systems Data Min. Knowl. Discov., 6 (1) (2002), pp. 83-105 in press
[25] Abdelaziz Berrado and George C. Runger. Using metarules to organize and group discovered association rules. Data Mining and Knowledge Discovery, 14(3):409–431, April 2007. in press
[26] Azmi, Mohamed, George C. Runger, et Abdelaziz Berrado. « Interpretable Regularized Class Association Rules Algorithm for Classification in a Categorical Data Space ». *Information Sciences* 483 (mai 2019): 313-31 in press
[27] Jerome H. Friedman & Nicolas I.Fisher. « Bump Hunting in High Dimensional Data », s. d. in press
[28] M. Holsheimer, M. Kersten, A. Siebes, Data surveyor: searching the nuggets in parallel, in: U.M. Fayyad, Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (Eds.), Advances in Knowledge Discovery and Data Mining, AAAI Press, Cambridge, 1996, pp. 447–467.
[29] Wolfgang Polonik, et Zailong Wang. « Prim analysis ». *Elsevier*, s. d
[30] Chong, I, et C Jun. « Flexible Patient Rule Induction Method for Optimizing Process Variables in Discrete Type ». *Expert Systems with Applications* 34, nº 4 (mai 2008): 3014-20. in press
[31] Wu, Longyang, et Hugh Chipman. « Bayesian Model-Assisted PRIM Algorithm », s. d., 6. in press
[32] Abu-Hanna, Ameen, Barry Nannings, Dave Dongelmans, et Arie Hasman. « PRIM versus CART in Subgroup Discovery: When Patience Is Harmful ». *Journal of Biomedical Informatics* 43, nº 5 (octobre 2010): 701-8, in press
[33] Díaz-Pachón, Daniel A., Jean-Eudes Dazard, et J. Sunil Rao. « Unsupervised Bump Hunting Using Principal Components ». *ArXiv:1409.8630 [Stat]*, 30 septembre 2014, in press
[34] Dazard, Jean-Eudes, et J. Sunil Rao. « Local Sparse Bump Hunting ». *Journal of Computational and Graphical Statistics* 19, nº 4 (janvier 2010): 900-929, in press
[35] Nannings, Barry, Ameen Abu-Hanna, et Evert de Jonge. « Applying PRIM (Patient Rule Induction Method) and Logistic Regression for Selecting High-Risk Subgroups in Very Elderly ICU Patients ». *International Journal of Medical Informatics* 77, nº 4 (avril 2008): 272-79, in press
[36] Charlie Pollack B, Ec, et . F.I.A.A. « A Comparison of PRIM and CART for Exploratory Analysis », s. d. in press

[37] Frikke-Schmidt, Ruth, Anne Tybjærg-Hansen, Peter Schnohr, Gorm B. Jensen, et Børge G. Nordestgaard. « Common Clinical Practice versus New PRIM Score in Predicting Coronary Heart Disease Risk ». *Atherosclerosis* 213, nº 2 (décembre 2010): 532-38. in press

[38] Sadiq, Saad, Yudong Tao, Yilin Yan, et Mei-Ling Shyu. « Mining Anomalies in Medicare Big Data Using Patient Rule Induction Method ». In *2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*, 185-92. Laguna Hills, CA, USA: IEEE, 2017.

[39] Maissae Haddouchi and Abdelaziz Berrado. 2018. Assessing interpretation capacity in Machine Learning: A critical review. In *Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications* (SITA'18). ACM, New York, NY, USA, Article 49, 6 pages. unpublished

Biographies

Dr. Abdelaziz Berrado is Department Chair and Associate Professor of Industrial Engineering in EMI School of Engineering at Mohamed V University in Rabat. He holds a Ph.D. degree in Decision Systems and Industrial Engineering from the Ira A. Fulton School of Engineering at Arizona State University. His research, teaching and consulting interests are in the areas of Big Data Analytics, Industrial Statistics, Operations and Supply Chain Modelling, Planning and Control with applications in healthcare, education and other industries. He focuses on developing frameworks, methods and tools for systems" diagnostics, optimization and control with the aim of operational excellence. He published several papers in research journals and conferences with local and international funding. He reviews for many journals and is member of INFORMS, IEOM and IEEE. In addition to academic work, he interacts continuously with different Industries in the areas of Machine Learning, Quality Engineering and Supply Chain Management. He was also a senior engineer at Intel.

Rym Nassih is a second year Phd student in Data Mining and Big data in EMI School of Engineering at Mohamed V University in Rabat. Her doctoral main research investigates data mining algorithms. She holds an engineering degree in business intelligence from the INSEA, National Institute of Statistics and Applied Economu in Rabat.