ISSN 2278-3091
Volume 8, No.1.4, 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse3981.42019.pdf
https://doi.org/10.30534/ijatcse/2019/3981.42019

Constraints Related to the Integration of the Information and Communication Technologies for Secondary Learners in the Learning of Physical Sciences in Morocco

OussamaDardary[1], Zineb Azar[1,2], Malika Tridane[1], Said Belaaouad[1]

[1]Laboratory of Physical Chemistry of Materials LPCM, Ben
M'Sik Faculty of Sciences. Hassan II University of Casablanca. Morocco.
dardaryosama@gmail.com, tridane.malika@gmail.com,
sbelaaouad@yahoo.fr
[2] Regional Center for Education and Training Casablanca
AnfaBdBirAnzarane Casablanca. Morocco.
azar.zineb@gmail.com

ABSTRACT

During their university studies, students encounter several problems and challenges which hinder their normal courses and which leads to dropping-out.
Our article will be dealing with a study of the problem of dropping-out at university. This study includes dynamic programming along with machine learning. We will focus on two main methods which are KNN, which based on the notion of similarity between data and the well-known method SVM, which can crack the problems of classification. We will also derive practical solutions using predictive analytics. And this would include application making predictions with real world example from University of Faculty of Chariaa of Fez.

Key words : Dynamic Programming, Machine Learning, Svm, Knn, Predictive Analytics.

1. INTRODUCTION

Thanks to the diverse nature of its higher education, Morocco is comparatively one step ahead compared to other countries in fighting the problem of dropping out of students. Undergraduates often leave higher education without a diploma. This reflects logistical and functional problems in our education system, and affects our society economically. We are, however, confident that we can reduce the number of dropping out. This objective can be achieved by the university, the government, researcher and experts. All of them can work together and put their complete efforts to combat this phenomenon. The Plan "Succeeding license" offers new opportunities for students to succeed, particularly the most fragile by assuring university autonomy, and with their increasing professionalism. [1]

Our research aims at realizing a system to give up dropping out at university. The challenge, nevertheless, is that we must be able to identify students are most likely to leave the university. This is achievable through a model, which is a table that includes all the characteristics of the students from a socio-economical perspective. This includes the students' age, sex, type of diploma, economical status - if they have a low economical status –and so on.

2. WHAT IS MACHINE LEARNING?

A. Definition Of Machine Learning

Machine learning refers to the development, analysis, and implementation of methods that allow a machine to evolve through a learning process and thus perform tasks that are difficult or impossible to achieve by more conventional algorithmic means.

Machine learning can be used for several purposes such as Speech recognition [2], Optical character recognition [3] and Recommender System [4]. A well-posed machine learning problem contains three main components:

- Performance measure to evaluate the learned system.
- Training experience to train the learning system.
- Types of problems and tasks.

B. How Machine Learning Relates To Statistics And Data Analysis?

The connection between the two fields: statistics and machine learning always poses an imposing question and if they should be separate domains or merge intimately. We can say that they are two complementary fields with different visions. And they have much in common since they both provide analytical methods. Statistics care about methods such as hypothesis testing, experimental design, ANOVA method, linear and logistic regression, generalized linear models, principal component analysis, factor analysis and discriminate analysis, to cite a few.

Machine learning cares about computational modeling and high-dimensional data. It includes advanced non-linear techniques, such as maximum margin methods, example support vector machines, SVMs, that uses special geometry in building an optimum prediction model. Other techniques include deep neural networks that have proven performance in several fields such as character recognition and machine translations. [4]

3. MACHINE LEARNING

A. Support Vector Machines (SVM)

The resolution of classification issues may be made by the SVM method inspired by the statistical learning theory; it was introduced by Vapnik in 1995. [5]

Support Vector Machines, or SVMs, belong to the family of maximum margin classifiers. In other words, it looks for the separating hyperplane (figure 1)with the largest margin with generalization guarantees. There are other advantages, including handling nonlinear problems using the kernel function to find nonlinear boundaries, and also allowing overlap in the classes.

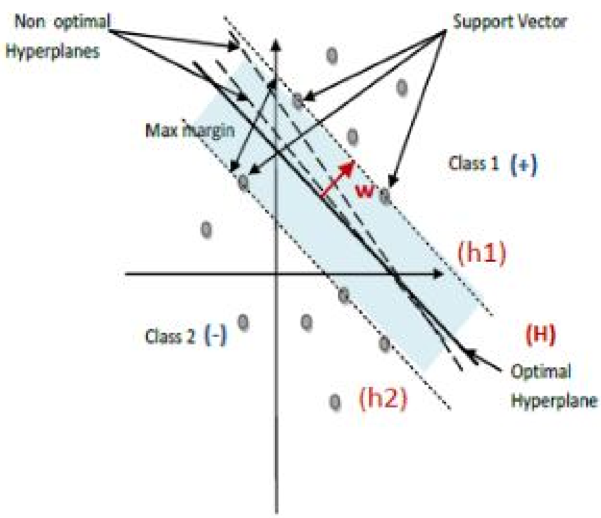

Figure 1: Hyperplane separating data belonging to two classes

The hyperplane H satisfies the equation [6]:

$$H: wx_i + b = 0 \quad (1)$$

With:
w The weight vector normal to the hyperplane.
b The bias or the intercept.
d+ The shortest distance from the hyperplane h1 to the closest positive example.
d- The shortest distance from the hyperplane h2 to the closest negative example.
With simple calculations using the definition of the distance between a point and the hyperplane, we have, as a result, that:

$$d_+ = d_- = 1/\|w\| \quad (2)$$

For instance, to be classified correctly, we need to satisfy one of these two constraints for each of h1 and h2, which we could combine together by multiplying by the label to make the constraint that we will include in the objective function.

$$w.x_i + b \geq +1 \; if \; y_i = +1 \quad (3)$$
$$w.x_i + b \leq -1 \; if \; y_i = -1 \quad (4)$$

The maximum margin classifier is the function that maximizes the geometric margin 1/||w ||.
The standard approach for binary problems is then to solve the soft margin formulation in which maximizing the margin, 1/||w ||.
SVMs have also the advantage to be a kernel method that allows generating non-linear decision boundaries.

To illustrate this notion, consider this two-dimensional non-linearly separable case in which only an ellipsoid (figure 2) can separate the data points in the original space x1 x2.[7]

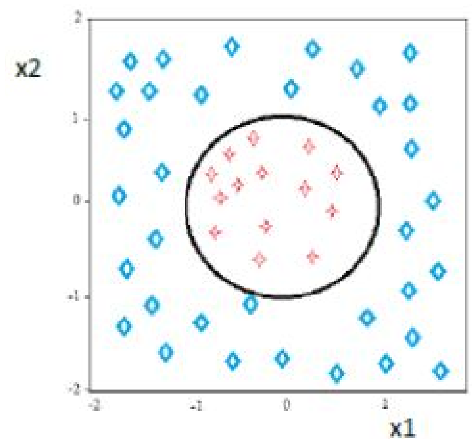

Figure 2: Ellipsoid separating the data points

However, with some mapping of x1 and x2 into a 3D space using two-degree polynomials (Figure 3), it is possible to transform the problem from a linearly separable problem into a linearly separable problem in a higher dimensional space.

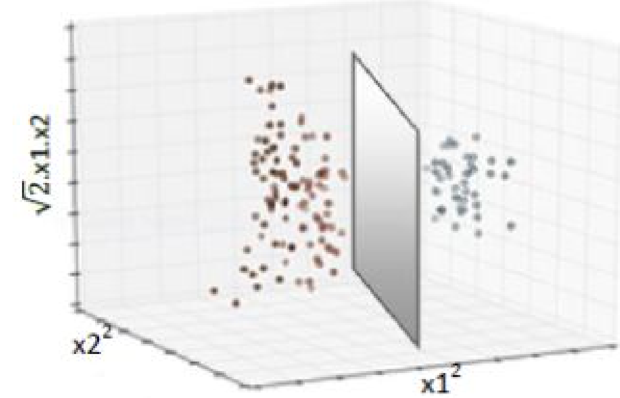

Figure 3: Linear separating in a 3D space

Technically, a kernel computes the dot product between the data points in higher dimensional space. [8]

B. K-Nearest-Neighbor(KNN)

KNN is a very simple and straightforward approach. Its principle is the following; unknown class data is compared to all stored data. We choose for the new data the majority class among its K nearest neighbors (It can therefore be heavy for large databases) in the sense of a chosen distance. The main idea of KNN is to use the notion of similarity between data.
In order to find the K Nearest-Neighbor to an example to be classified, we can choose the Euclidean distance.

Considering two examples represented by two vectors xi and xj, the distance between these two examples is given by[9]:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{d}(x_{ik} - x_{jk})^2} \quad (5)$$

4. THE PROBLEM OF UNIVERSITY DROPOUTS

“The fact, for a student, to leave higher education without obtaining a diploma” This problem has negative effects on students’ situation, especially when they hope for getting their degrees. And the most negative effect is the psychological one. [10]

Drop-out problem is also a serious issue for higher education institutions. Abandonment rates inevitably affect the quality of training, its organization, its production and its contents. [11][12]

Finally, dropout problem is an economical one as well. This involves the society and the government that invest heavily in the expenditure of higher education. [13]

A. Data Description & Preparation

This part describes the data we use in this automatic learning application. We obtained the diffusion data for the study of the prediction of the university dropout of the Faculty of Chariaa between 2011 and 2012.

This is a statistical study of approximately 3000 students who are subscribe from 2011-2012. The frequency of university drop-out, as shown in this table, is about 30%.

Table 1: Number of student drop-out

University drop-out	Cases
Indicated before the end of the first year	909
Indicated after the first year	315

We constructed a large matrix of about 3000 students described by 86 variables. Overall, you can see, in Fig. 4, a mixture of binary functions, numeric functions, categorical characteristics, and a large amount of missing values in the data.

B. Complexity Of Data

We have prepared this data to feed it to machine learning algorithms. Positive examples, drop-out students, are disseminated in space, which poses a challenge to find a global predictive model discriminating positive from negative examples.

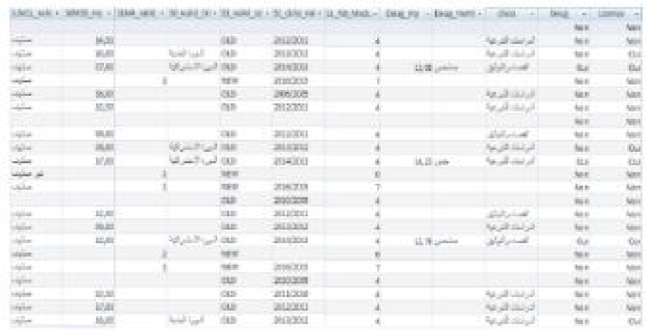

Figure 4: Snapshot of described variables of students

First, we handle the complexity of data by organizing functionalities into groups. For example, DMG, a demographic group, includes characteristics such as student age, marital status, family support, income - an indicator of the student's socio-economic status, gender, Etc.

The second group, PSH, stands for Previous Study History. We also have information on the status of modules, SM, that includes both normal and catch-up sessions for each semester (S1, S2, S3, S4, S5 and S6). We also have all the information concerning the type and the grade of the baccalaureate and the region where it was taken (IB).

The calendar of the data is illustrated in this figure (Fig. 5) At each step, S1, S2, S3, S4, S5 and S6, a set of characteristic groups is collected. Thus, we have these different temporal points, T0, T1, T2, T3, T4 and T5 which represent semester one, semester two, semester three, semester four, semester five, and semester six respectively. In each semester, there is a group of characteristics that were collected– which we described earlier.

Also, by separating the data into multiple data sets, we can focus on specific sub-tasks of university dropouts in the effort to devise a refined model.

Table 2: Number of University Drop-Out in Different Temporal Points

	T0	T1	T2	T3	T4	T5
features	10	14	14	12	12	12
points	909	71	62	19	113	50

C. Methods *For Prediction Of This Problem*

In this study, we come to the University drop-out prediction as a binary classification problem. Students who abused the studies are assigned the positive class, while those who completed the studies are assigned to the negative class. Each Student is a data point, also called example, and described by feature vector x and the discrete label y. y could be -1 or +1. It could be also 0 or 1.

classification function f. Once we have f, we can use it to classify a new student. [14]

Our concrete aim is to obtain a model with low test errors.

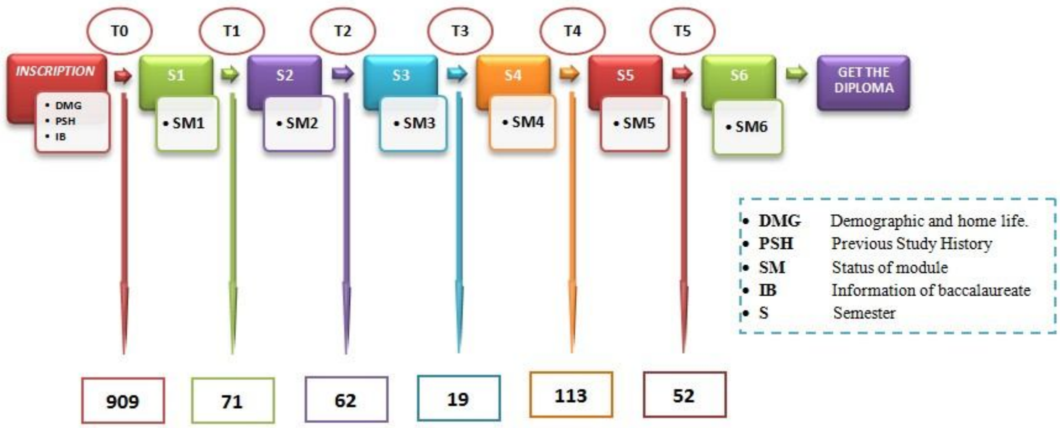

Figure 5: Calendar of the characteristic collected

We would like to predict at different time points, and hence, we have training data sets for each take T0, T1, T2, T3, T4 and T5. We applied the simple nearest neighbor algorithm to the data.

In Table 4, we count the average number of positive points in the neighborhood of each positive point. For the distance, in general, choosing the distance metric is a hard problem. And the ideal is to learn one from the data. But to simplify, we chose the Euclidian distance. We use different values of k, and we illustrate here with k equal 12. Results show that the university dropout cases, or the positive examples, have between 0 and 4 positive examples in their neighborhoods. In other words, Positive examples, which represent the minority class, are isolated among negative examples. We expected, then, that this phenomenon would lead to challenges in finding a general model that fits the data.

Table 3: Data Distribution

	Euclidian
T0	3,63 1,9
T1	3,45 1,72
T2	2,61 2,48
T3	4,04 2
T4	3,37 2,37
T5	5,03 2,03

We cast the problem as a classification problem, seeking to use data to derive a binary classification function $f: \mathbb{R}^d \rightarrow \{-1, +1\}$. The general picture is to use the data at hand as input to train machine learning algorithms to find the

We first build the confusion matrix to report the misclassification errors [15].

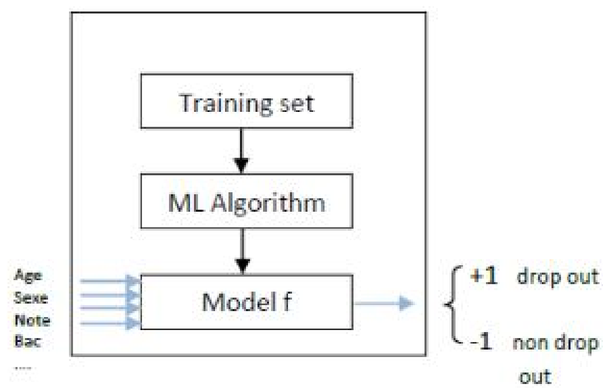

Figure 6: Formalization Problem

Table 4: Confusion Matrix

		Actual Label	
		UDO	Non UDO
Predicted Label	UDO	Positive(P)	Non Positive(NP)
	Non UDO	Non Negative(NN)	Negative(N)

We first build the confusion matrix to report the misclassification errors.

UDO University dropout.

P Dropping out students who were predicted as university dropout by the model.

N Non dropping out students who were predicted as non university dropout by the model.

NP Non dropping out students who were predicted as university dropout by the model.

NN Dropping out students who were predicted as non university dropout by the model.

$$Sensitivity = \frac{P}{P+NN} \quad (6)$$

$$Specifity = \frac{N}{N+NP} \quad (7)$$

Sensitivity is the probability that a test performed on a dropout student is positive; in other words, the test is positive knowing that the student has dropped out.

Specificity is the probability that a test performed on a non dropout student is negative; in other words, the test is negative knowing that the student has not dropped out.

The negative class constitutes the majority. Therefore, it is not difficult when we create a module for the goal of obtaining high specificity rates, with a click, to insure a reasonable balance amid specificity and sensitivity. Keep in mind that the motive for choosing these two measures is the purpose of being able to accurately forecast students with the uppermost risk for leaving university, in order to initiate interventions, we use geometric mean, or G, which is [16]:

$$G = \sqrt{Sensitivity * Specifity} \quad (11)$$

5. RESULTS AND DISCUSSION

In order to assess our findings, the next process is to be repeated: a train and test set of 80 to 20 ratio constitute a random partition of every data set. A test set is only used to assess the model or analyze the out of sample error. Therefore, it is never used during a learning phase. Each class is split equivalently amid the sets. After that, in order to elect the best model and ideal parameters, we put into application a five-fold cross validation to the training set. Finally, the elected model can be tested on the hidden test set, and the recording process of confusion matrices starts off for different subsets of data alongside sensitivity, specificity, and the G for every single one.

Table 5: Results Table

	Sensitivity	Specificity	G
T0	0,41	0,67	0,52
T1	0,75	0,96	0,85
T2	0,79	0,89	0,98
T3	0,75	0,98	0,86
T4	0,86	0,97	0,91
T5	0,73	0,82	0,77

After studying an approximate number of 3,000 students from the Faculty of Chariaa of Fez, we can conclude that the SVM is a realistic and reasonable method that has allowed us to achieve results close to reality.

Note that some variables can change over time (family situation, economic ...) are influencing the results. It can also be seen that in T4 and T5 the number of students decreases remarkably, this big fall can have an explanation that some students who got their DEUG diploma could have either applied for other jobs or have enrolled in other universities or institutions.

6. CONCLUSION

This confirms that a more exact forecast of dropping out is not an evasive task. However, this is not the end. Stronger models with advanced estimate capacity ought to be built in order to classify students that are most probably dropping out of university. The borderlines of machine learning are challenged due to the dropout problem. This is due to the challenging nature of the latter, which is an everyday problem in our world. This issue is challenging for a number of factors. One factor is the missing data while a second one is the temporal nature of dropping out.

Our work opens the door on a new perspective that takes into account other problems such as the guidance of the students, the relationship between the university and the employment.

REFERENCES

[1] Thierry Côme , Abdelilah Yassine Accès Régulé A L'université Et Implication Et Motivation Des Etudiants : L'exemple Du Maroc Gestion Et Management Public, 2015/2 (Volume 3 / N° 4) https://doi.org/10.3917/gmp.034.0005

[2] Tomi Kinnunen, Haizhou Li. An Overview Of Text-Independent Speaker Recognition: From Features To Supervectors. Speech Communication 52 (2010) 12–40,2009. https://doi.org/10.1016/j.specom.2009.08.009

[3] Shunji Mori, Hirobumi Ishida Hiromitsu Amada, Optical Character Recognition 1st John Wiley & Sons, Inc. New York, Ny, Usa ©1999 Isbn: 0471308196.

[4] Bruce Ratner, Statistical And Machine-Learning Data Mining: Techniques For Better Predictive Modeling And Analysis Of Big Data, Second Edition, Isbn 9781439860915 - Cat# K12803

[5] Vapnik, V. Machine Learning (1995) 20: 273. doi: 10.1023/A : 1022627411411

[6] Christina Leslie, Eleazar Eskin, William Stafford Noble, The Spectrum Kernel: A String Kernel For Svm Protein Classification, Pacific Symposium On Biocomputing 7:566-575 (2002).

[7] M. Premalatha1 , C. Vijaya Lakshmi2, Svm Trade-Off Between Maximize The Margin And Minimize The Variables Used For Regression International Journal Of Pure And Applied Mathematics Volume 87 No. 6 2013, 741-750 Issn: 1311-8080 (Printed Version); Issn: 1314-3395 (On-Line Version) https://doi.org/10.12732/ijpam.v87i6.2

[8] Everything You Wanted To Know About The Kernel Trick , Eric Kim, Http://Www.Eric-Kim.Net

[9] Ramaswamy, S.; Rastogi, R.; Shim, K. (2000). Efficient Algorithms For Mining Outliers From Large Data Sets. Proceedings Of The 2000 Acm Sigmod International Conference On Management Of Data – Sigmod '00. P. 427. Doi:10.1145/342009.335437. Isbn 1-58113-217-4.

[10] François Fillon, Réduire De Moitié Le Décrochage Universitaire, Juin 2011.

[11] (Texte Présentant De L'intérêt Pour L'eee) Concernant Les Politiques De Réduction De L'abandon Scolaire,2011

[12] Diagnostic Stratégique Du Partenariat Institutionnel,Etude Du Conseil Superieur De L’enseignement ; 2009

[13] Simon Field, Małgorzata Kuczera, Beatriz Pont, En Finir Avec L’échec Scolaire Dix Mesures Pour Une Education Equitable, Les Corrigenda Des Publications De L’ocde Sont Disponibles Sur : Www.Oecd.Org/Editions/Corrigenda. © Ocde 2007 https://doi.org/10.1787/9789264032620-fr

[14] A Practical Guide To Support Vector Classification Chih-Wei Hsu, Chih-Chung Chang, And Chih-Jen Lin Department Of Computer Science National Taiwan University, Taipei 106, Taiwan .Version: 2003 Last Updated: May 19, 2016

[15] Simple Guide To Confusion Matrix Terminology Http://Www.Dataschool.Io/Simple-Guide-To-Confusion-Matrix-Terminology/

[16] Ilia Vovsha , Ashwath Rajan , Ansaf Salleb-Aouissi , Anita Raja, Axinia Radeva , Hatim Diab , Ashish Tomar and Ronald Wapner Predicting Preterm Birth Is