A logo for university computing

Auto-generated description

**Evaluation cover page**

|  |  |
| --- | --- |
| *Student's Full Name:* | Rosemary Dejesus Ramirez Cords. |
| *Student Number:* | 2024017. |
| *Module Title:* | Machine Learning. |
| *Evaluation Title:* | CA1 Project |
| *Assessment Due Date:* | 21/04/2024. |
| *Presentation Date* | 21/04/2024. |

**Declaration**

By submitting this review, I confirm that I have read CCT's policy on academic misconduct and understand the implications of submitting work that is not mine or that does not appropriately reference material taken from a third party or other source.

I declare that it is my own work and that all third-party material has been properly referenced.

I further confirm that this work has not previously been submitted for evaluation by me or anyone else at CCT College Dublin or any other higher education institution.

**CA1 – Project**

**Students Adaptability Level in online Education**

[**Introduction** 2](#_Toc164151633)

[**Characterization of data and pre-processing** 2](#_Toc164151634)

[**Training and Testing our Data Set** 2](#_Toc164151635)

[**Scaling and Normalization** 2](#_Toc164151636)

[**Applying Modals** 3](#_Toc164151637)

[**Hyperparameter** 10](#_Toc164151638)

[**Conclusion** 11](#_Toc164151639)

[**References** 11](#_Toc164151640)

**CA1 – Project**

**Students Adaptability Level in online Education**

# **Introduction**

Technology that has penetrated every aspect of our life has altered teaching and learning. Internet and Communication Technology (ICT) has rewritten the rules. The university is no longer the sole repository of scholarship. The class is not enclosed within the walls of the classroom anymore. Knowledge is not contained in a textbook. Imparting it is not the domain of the teacher. The degree is not the sole proof of learning. Education, as we have known it, is on the cusp of a profound change. Gutenberg’s printing press made books easier to print, and what had been handwritten, rare, precious, and so tied to library shelves was freed of the chains. The computer and internet gave us the ‘soft copy’ that freed information from all physical media. ICT is virtually opening education to the whole world. (Harish Janani,2013)

Going through the present data we display independant variables (Gender, Age, Education Level, Institution Type, IT Student, Location in Town, Load-shedding, Financial Condition, Internet Type, , Network Type Class Duration, Self LMS and Device), and dependant variable (Adativity Level). The goal is to approximate the mapping function so well that when we have new input (independant variables) we can predict the output variables (dependant variables).

In short words, we will apply Supervised Machine Learning modals in regards to predict their adaptability in three classifications (Moderate, Low or High) based on certain characteristics of students.

# **Characterization of data and pre-processing**

Exploratory Data Analysis or (EDA) is understanding the data set by summarizing its main characteristics and often plotting them visually. This step is very important especially when we arrive at modelling the data to apply Machine learning. Plotting in EDA consists of Histograms, Box plot, Scatter plots and many more. Through the process of EDA, we can also refine the problem statement or definition of our problem.(McQuaid, D. (2024b). file:///C:/Users/Dell/Downloads/Feature%20Scaling%20or%20Normalization%20(3).pdf.).

Also, define the characteristics of our data (number of columns, rows, null values, etc). In the following part we will display it.

This data is a csv document which we import as the name of “df” in our Jupyter Notebook.

A white rectangular object with a white line

Description automatically generated

**Fig1: Display library to import data set.**

In order to know how many columns, rows and which data types we have we display the function .info

A screenshot of a computer

Description automatically generated

**Fig2: Display rows, columns and data types.**

In this case our data have 1205 rows (observations) and 14 columns (features) with object as a data type which means string values.

In regards to know if we have null values in our data set, we use the function .isnull().sum()

A screenshot of a computer

Description automatically generated

**Fig3: Display null values.**

The function show that our data set don’t have any null values.

We use histograms to compare our target (Adaptivity Level) with our different features.

A screenshot of a computer screen

Description automatically generated

**Fig 4: Display target versus features.**

When we want to apply our modals, we need to replace categorical values for numerical values.

A screenshot of a computer

Description automatically generated

**Fig 6 : Replace categorical values for numerical values.**

In this case the components of our target are Moderate, Low and High which we replace for 0,1 and 2 respectively.

Also, we need to define our dependant(Y) and independent(X) values, with function .shape we can see how many columns and rows we have. After that we take from column 0 to 13 our independent values and column 13 for dependant values. Remember that our modal counts the column from 0.

A screenshot of a computer program

Description automatically generated

**Fig 7: Define values for dependant values and independent values.**

So, finally in the function .shape display the shape for X and Y

How we transform our target in numerical values we need to transform our features as well to apply the models. In this case we use the function .dummies to do it.

A screenshot of a computer

Description automatically generated

**Fig 8: Replace categorical values for numerical values.**

# **Training and Testing our Data Set**

Regarding apply the model we need to split the data for train and test. In this first case we use for test size 25% .

**A screenshot of a computer code

Description automatically generated**

**Fig 9: Splitting data.**

For test size of 25% we will use 302 observations and 35 features.

# **Scaling and Normalization**

Some features, such as latitude or longitude, are bounded in value. Other numeric features, such as counts, may increase without bound. Models that are smooth functions of the input, such as linear regression, logistic regression, or anything that involves a matrix, are affected by the scale of the input. Tree-based models, on the other hand, couldn’t care less. If your model is sensitive to the scale of input features, feature scaling could help. As the name suggests, feature scaling changes the scale of the feature. Sometimes people also call it feature normalization. Feature scaling is usually done individually to each feature. Next, we will discuss several types of common scaling operations, each resulting in a different distribution of feature values.( McQuaid, D.(2024a)file:///C:/Users/Dell/Downloads/Feature%20Scaling%20or%20Normalization%20(3).pdf.)

So, before to apply our models, normalization is necessary to perform our points of data and have the same measure on it.

A screenshot of a computer

Description automatically generated

**Fig 10: Standard Normalization.**

In this case we use Standard Normalization regarding have the same measure and apply the modals.

# **Applying Modals**

Machine learning algorithms that learn from input/output pairs are called supervised learning algorithms because a “teacher” provides supervision to the algorithms in the form of the desired outputs for each example that they learn from.( Müller, A.C. and Guido, S. (2016). https://www.nrigroupindia.com/ebook/Introduction%20to%20Machine%20Learning%20with%20Python%20(%20PDFDrive.com%20)-min.pdf. Available at: <http://safaribooksonline.com/>.)

We apply fourth different Supervised Machine Learning Modals in this data set with three different training and testing 20%, 25% and 30%. This different testing allow us to compare our models with different splitting and their accuracy in our investigation.

I decided to use Artificial Network Intelligence, Decision Tree and Support Vector Machine because I have a classification target (Moderate, Low and High), in short words I will classified them.

**Artificial Network Intelligence (Deep Learning)**

Multilayer perceptrons (MLPs) are also known as (vanilla) feed-forward neural networks, or sometimes just neural networks. MLPs can be viewed as generalizations of linear models that perform multiple stages of processing to come to a decision.( Müller, A.C. and Guido, S. (2016). https://www.nrigroupindia.com/ebook/Introduction%20to%20Machine%20Learning%20with%20Python%20(%20PDFDrive.com%20)-min.pdf. Available at: <http://safaribooksonline.com/>.)

I didn’t use cross validation for this model because didn’t support it.

**Testing 25%**

A screenshot of a computer program

Description automatically generated

**Fig. 11: Score of training and testing.**

A screenshot of a graph

Description automatically generated

**Fig.12: Prediction across the three classes by the confusion matrix with accuracy.**

**Testing 20%**

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a graph

Description automatically generated**

**Fig.13: Prediction across the three classes by the confusion matrix with accuracy.**

**Testing 30%**

**A screenshot of a computer program

Description automatically generated**

**Fig. 14: Accuracy of model in training and testing.**

**A screenshot of a graph

Description automatically generated**

**Fig.15: Creation of predictions across the three classes by confusion matrix.**

**Decision Tree**

Decision trees are widely used models for classification and regression tasks. Essen‐ tially, they learn a hierarchy of if/else questions, leading to a decision.( Müller, A.C. and Guido, S. (2016). https://www.nrigroupindia.com/ebook/Introduction%20to%20Machine%20Learning%20with%20Python%20(%20PDFDrive.com%20)-min.pdf. Available at: <http://safaribooksonline.com/>.)

**Testing 25%**

**A screenshot of a computer

Description automatically generated**

**Fig.16: Predictions of three classes by confusion matrix with accuracy and recall report. Also after apply the modal we want to be sure about the prediction across the all folds classifier so, we apply cross validation.**

**Testing 20%**

**A screenshot of a computer program

Description automatically generated**

**Fig.17: Predictions of three classes by confusion matrix with accuracy and recall report. Also after apply the modal we want to be sure about the prediction across the all folds classifier so, we apply cross validation.**

**Testing 30%**

A screenshot of a computer

Description automatically generated  
**Fig.18: Predictions of three classes by confusion matrix with accuracy and recall report. Also after apply the modal we want to be sure about the prediction across the all folds classifier so, we apply cross validation.**

**Support Vector Machine.**

The two most common linear classification algorithms are logistic regression, imple‐ mented in linear\_model.LogisticRegression, and linear support vector machines (linear SVMs), implemented in svm.LinearSVC (SVC stands for support vector classi‐ fier). Despite its name, LogisticRegression is a classification algorithm and not a regression algorithm, and it should not be confused with LinearRegression.( Müller, A.C. and Guido, S. (2016). https://www.nrigroupindia.com/ebook/Introduction%20to%20Machine%20Learning%20with%20Python%20(%20PDFDrive.com%20)-min.pdf. Available at: <http://safaribooksonline.com/>.)

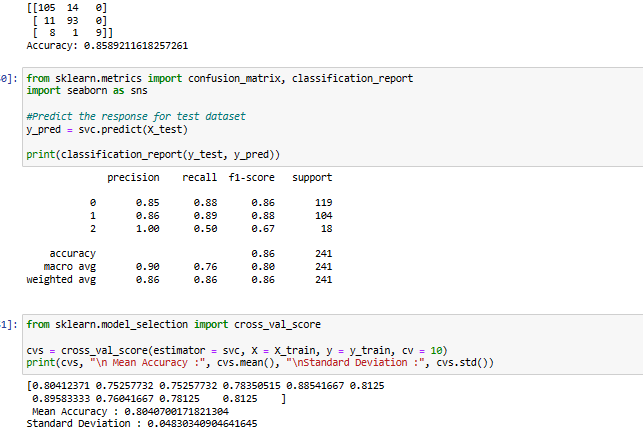
**Testing 25%**

**A screenshot of a computer

Description automatically generated**

**Fig.19: Predictions of three classes by confusion matrix with accuracy and recall report. Also after apply the modal we want to be sure about the prediction across the all folds classifier so, we apply cross validation.**

**Testing 20%**

****

**Fig.20: Predictions of three classes by confusion matrix with accuracy and recall report. Also after apply the modal we want to be sure about the prediction across the all folds classifier so, we apply cross validation.**

**Testing 30****%**

**A screenshot of a computer

Description automatically generated**

**Fig.21: Predictions of three classes by confusion matrix with accuracy and recall report. Also after apply the modal we want to be sure about the prediction across the all folds classifier so, we apply cross validation.**

**Compare modals 25% testing.**

**A screenshot of a computer program

Description automatically generated**

**Fig. 22: Comparison between the three models.**

**Compare modals 20% testing.**

**A screenshot of a computer program

Description automatically generated**

**Fig. 23: Comparison between the three models.**

**Compare modals 30% testing.**

**A screenshot of a computer program

Description automatically generated**

**Fig. 24: Comparison between the three models.**

# **Hyperparameter**

The primary purpose of hyperparameter tuning is to tweak model performance for optimal results.  
According to a very popular book called “Applied Predictive Modelling” - “Many models have important parameters which cannot be directly estimated from the data. For example, in the K-nearest neighbor classification model … This type of model parameter is referred to as a tuning parameter *because there is no analytical formula available to calculate an appropriate value.*”

We could elaborate on specific hyperparameter tuning techniques applied to machine learning models to find optimal parameters, for example:

* The learning rate for training a neural network.
* The C and sigma hyperparameters for support vector machine.
* The K in k-nearest neighbors.

To demonstrate how the hyperparameters works, I decided to apply in the Support Vector Machine Model where with the 30% of testing I got 80% accuracy.

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

A close-up of a computer screen

Description automatically generated

**Fig. 25: New accuracy, recall and predictions across the three classes by confusion matrix.**

How you can see on the top, I got a better accuracy(91%) with a better distribution in the confusion matrix.

# **Conclusion**

Was useful for this investigation to use the Supervised Machine Learning algorithms because they allow us to made predictions about the adaptability of students in online education bellow three different categories (Moderate, Low and High).

Between the three of them I consider the most accurate Model was the Decision Tree Model with the 30% of testing because how I demonstrated in the report, it creates a good confusion matrix making predictions between the three classes, from 362 samples, classified correctly for class 0, 169, class 1, 135 and class 2, 24, with 91% of accuracy and very close recall which is important in regard to trust our models and predictions. Also applying the cross validation reducing the risk of underfitting or overfitting our classifier (a technique for estimating the performance of a predictive model across different folds of data) display a mean accuracy approximately of 0.887 with a standard deviation of about 0.045, which means that the model classifier performs well, but there is some variability in its performance across different folds.

Also, after use a hyperparameter in regard to tuning our modal should be useful apply the Cross Validation because we are receiving a better accuracy but is not less important be sure about what happens with the predictions in different fold of data.

# 

# **References**

* Harish, Janani. Online Education: A Revolution in the Making. Vol. 2, no. 1, 2013, pp. 26–38, www.cadmusjournal.org/files/pdfreprints/vol2issue1/reprint-cj-v2-i1-online-education-jharish.pdf. Accessed 11 Apr. 2024.
* McKinney, W. (2018). *Python for Data Analysis Data Wrangling with Pandas, NumPy, and IPython*. [online] file:///C:/Users/Dell/Downloads/Python%20for%20Data%20Analysis.%20Data%20Wrangling%20with%20Pandas,%20NumPy,%20and%20IPython%20(2017,%20O%E2%80%99Reilly)%20(3).pdf. Available at: <http://oreilly.com/catalog/errata.csp?isbn=9781491957660>.
* McQuaid, D. (2024a). *Feature Scaling or Normalization*. file:///C:/Users/Dell/Downloads/Feature%20Scaling%20or%20Normalization%20(3).pdf.
* McQuaid, D. (2024b). *What is Exploratory Data Analysis?* file:///C:/Users/Dell/Downloads/Feature%20Scaling%20or%20Normalization%20(3).pdf.
* Müller, A.C. and Guido, S. (2016). *Introduction to Machine Learning with Python A GUIDE FOR DATA SCIENTISTS*. [online] https://www.nrigroupindia.com/e-book/Introduction%20to%20Machine%20Learning%20with%20Python%20(%20PDFDrive.com%20)-min.pdf. Available at: <http://safaribooksonline.com/>.