

|  |
| --- |
| Project Cover Sheet |

|  |  |
| --- | --- |
| **TO BE FILLED BY THE STUDENT** | |
| Student Name:  Mohamed Zaki |  |
| Email ID: mzakiibrahim20@gmail.com |  |
| Date Submitted:  09/07/2023 |  |

**ASSESSMENT FEEDBACK**

|  |  |  |
| --- | --- | --- |
| **TO BE FILLED BY THE ASSESSOR** | | |
| Assessment type | Marks | Marks Awarded |
| Project Documentation | 50 |  |
| Program File | 50 |  |
| Overall Marks achieved |  | |
| GRADE ACHIEVED |  | |
| **Summative Feedback by Assessor for further improvement** | | |
|  | | |
| **Comments for REDO submission (If applicable)** | | |
|  | | |

**GRADE DESCRIPTORS**

|  |  |
| --- | --- |
| 70% and above  (Distinction) | The Project evaluated is of a high to exemplary standard. The work addresses clearly and articulately the project requirements and thus meets and satisfies all the learning outcomes (either well or in an exemplary way). The work demonstrates: clear knowledge; references to appropriate academic literature; analysis; critical evaluation; and originality of argument. It is structured and presented to a high (or exemplary) standard. Referencing conventions are fully observed. |
| 60 to 69%  (Merit) | The project evaluated is of a good to a high standard. Substantial knowledge, comprehension and analysis is evident throughout. Arguments presented are clear and focussed with a logical structure in place. There is clear evidence of critical evaluation of a wide range of theories/perspectives from academic literature and some independent thought. The work is well-written and addresses well all of the learning outcomes. Referencing conventions are fully observed. |
| 50 to 59%  (Pass) | The project evaluated is of a fair to good standard. Adequate knowledge, comprehension and analysis is evident throughout. The arguments presented have a logical structure and show some critical evaluation in places, although there may be limited evidence of an independent perspective. There is evidence of some good engagement with some of the appropriate literature. Learning outcomes have been largely met and to an appropriate degree. Referencing conventions are observed. |
| 40 to 49%  (Fail/Redo) | The project evaluated is of a basic standard. The arguments presented have some logical structure and are supported by academic literature in most cases. The academic literature used is outside of the suggestions made in the module guide but remains limited. Little critical evaluation is evident, and the work tends more widely towards a descriptive style. Learning outcomes have been addressed in a basic but satisfactory way. Referencing conventions are mostly observed. |
| Fail Grades | |
| 30 to 39%  (Module retake) | The project evaluated is of a limited standard. Limited use of academic literature and as such knowledge and argument is very weak. A simple descriptive style with no evidence of critical evaluation throughout. Over-reliance on simplistic, limited sources. Referencing conventions may not be observed. Some learning outcomes met but in a weak and simplistic way. The work is needs to be developed in greater depth and detail to move to a passable standard at this level of study |
| 29% and Below  (Module retake) | The project evaluated is of an unacceptable standard. There is little or no evidence of knowledge and understanding that is required at this level. Referencing is inadequate or non-existent. The learning outcomes have not been addressed fully and the work requires significant modification to bring it to a passable standard. |

**General Guidelines**

1. Complete the ‘To be filled by the student section’ in the cover page.: **Done**
2. Project documentation must be submitted as an electronic document in MS word (Use 12 Times New Roman script with 1.5 spacing between lines): **Done**
3. Submit the Project in MS word document with the file name being:

First Name Last Name \_ Program Code

1. Example: John Smith\_DS (DS abbreviated as Data Science) **Done**

**Advance Programs Capstone Project**

**Overview:**

In this capstone project, students will be required to demonstrate the application of data science and analytics skills learned throughout the course. The project involves the end-to-end process of addressing a real-world problem using data analytics. This includes data gathering, cleaning, exploratory data analysis, visualization, and applying machine learning algorithms for predictive analysis.

**Objectives:**

* Formulate a real-world problem that can be addressed using data analytics.
* Gather and clean data relevant to the problem.
* Conduct exploratory data analysis to understand trends, patterns, and anomalies in the data.
* Visualize data in a way that is meaningful and helps in decision-making.
* Apply appropriate machine learning algorithms to make predictions or draw conclusions.
* Evaluate the effectiveness of your solution.
* Communicate your results effectively to both technical and non-technical audiences.

**Sample Topics:**

1. Predictive Analysis in Retail: Predict future sales based on historical sales data for a retail company. Use exploratory data analysis to identify key sales trends and seasonality. Apply machine learning algorithms to predict future sales.
2. Customer Churn Analysis in Telecommunication: Analyze customer data to identify key factors contributing to customer churn. Develop a predictive model to identify customers at risk of churning.
3. Fraud Detection in Financial Transactions: Develop a predictive model that can identify fraudulent transactions based on historical transaction data.
4. Healthcare Diagnostics: Develop a predictive model that uses patient data (like age, gender, medical history, etc.) to predict the likelihood of a particular disease.

These are just examples, and you are encouraged to develop your own unique project related to your field of interest.

**Documentation:**

The following instructions provide guidance on how to compile the project documentation.

**Abstract**

The abstract must be a concise summary, encapsulating the project title, the methodology utilized, key findings, and the implications of your project. Try to maintain a word limit of 150-250 words for clarity and brevity.

‘’ this project objective is to provide and apply the learned Skills during the journey of Python learning in Data science , during the last four modules in Learning python Data Science , we have learned introduction to python , Data set types , how to clean and pre process the data sets and finally how to use AI and ML algorithms to predict output or model , with this project we will apply what we have learned in last months ‘’

**Introduction**

The introduction sets the context of your project. This should clearly state the problem you're addressing, its significance, and its relevance. Objectives should be well-defined, measurable, and directly related to the problem statement. It is also necessary to delineate the scope and any constraints associated with your project.

**‘’Project Title : Customer prediction for Churn**

In this project , I had downloaded a dataset from website **(**[Bank Customers Churn | Kaggle](https://www.kaggle.com/datasets/santoshd3/bank-customers)) for a bank dataset for all customers with different features , Age ,Gender, Geographic location , Balance ,….

And one of the features is one variable to define if customers Exit or not. the objective of this project is to apply different machine learning models to predict for any other customer with different Features value and define whether it would Exit or not.

The data set consists of Ten thousands Rows and fourteen columns, where Two thousand and thirty seven customers exited and seven thousands and nine hundreds and sixty three customers still exist .’’

**Methodology**

In this section, detail your research design, data collection and analysis methods, and your plan for executing the project. The methodology should be transparent and sufficiently detailed, enabling another researcher to replicate your study based on your description.

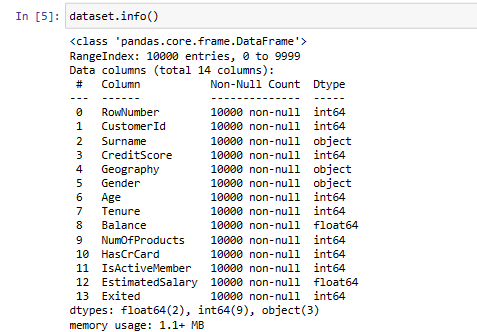
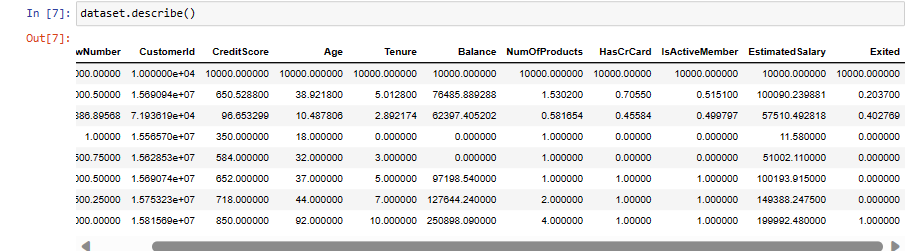
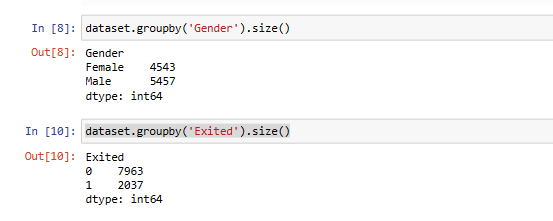
‘’From Dataset structure , first step is to define the type of dataset and what exactly you need to extract from this dataset , the dataset has features which represent the customer information such as Age, Geographic locations , Estimated Salary and Balance and other features which can be considered as in-dependent variables (Predictors) and there is on feature represent the status of customer , this feature can be considered as dependent variable . so in this dataset, we can supervised data learning as it has a clear target and also labeled data. but as we have a class in the target, means we would classify each customer either he will Exit or no. so we have to use supervised classification Machine learning algorithms such as Decision Tree, Random Forest , SVM ,…etc. .

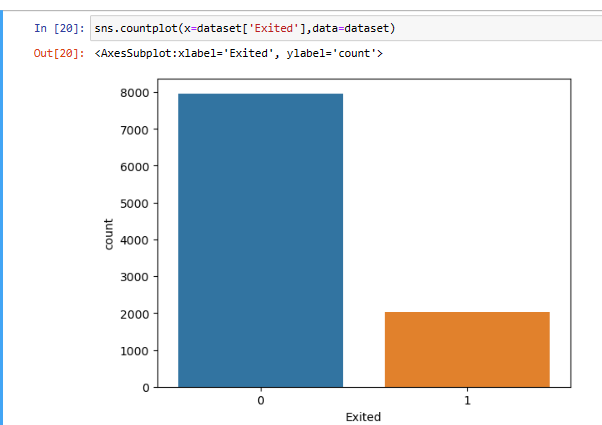
But before applying these algorithms, we should preprocess and clean the data set and make the necessary transformation to make it fit for module. ‘’

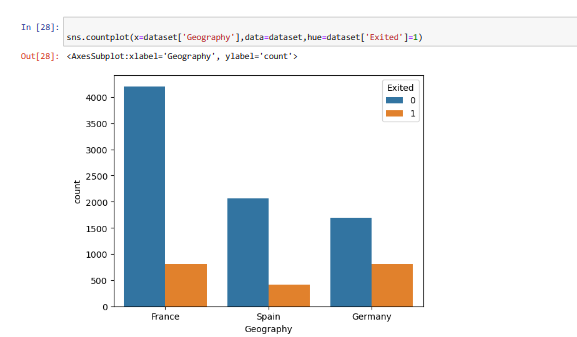
**Project Implementation**

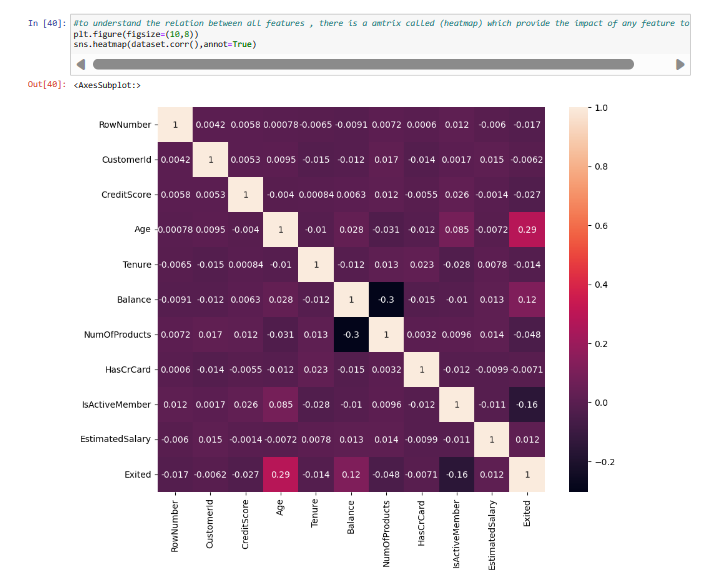
This section should provide a thorough account of your project's execution. Discuss your design considerations, the development process, and how you conducted any testing or validation of your work. Ensure the description is comprehensive and methodical.

‘’Starting by importing the datasets and have more information and description in dataset .

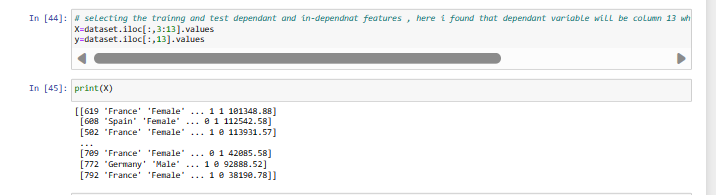
 

**t**hen do Data visualization to understand the relation between Features .



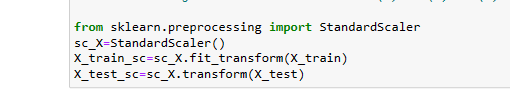


**Define Dependent and in-dependent variables:**

****

****

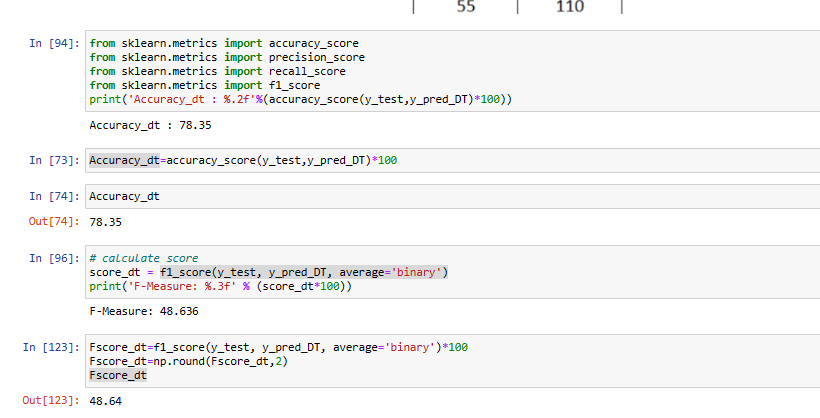
**Also in the course of Transformation , I have standard scaler to unify Feature values**

****

**Then splitting the dataset to testing and training data with test size of 0.20 and apply the machine learning algorithms.**

**And apply the performance parameters on each module and confusion matrix , here I have used the Accuracy Score and F1-Score for each module as per below sample .**

****

****

**Results and Discussion**

Here, present the outcomes of your project and engage in a critical discussion about their significance. Compare your results with previous work, talk about any unexpected findings, and speculate on the potential implications of your results.

Add output pictures and code snippets wherever its required to support your answer.

**Conclusion**

The conclusion should recap your project, emphasizing the main findings and their implications. Propose potential future research directions that could stem from your work.

**References**

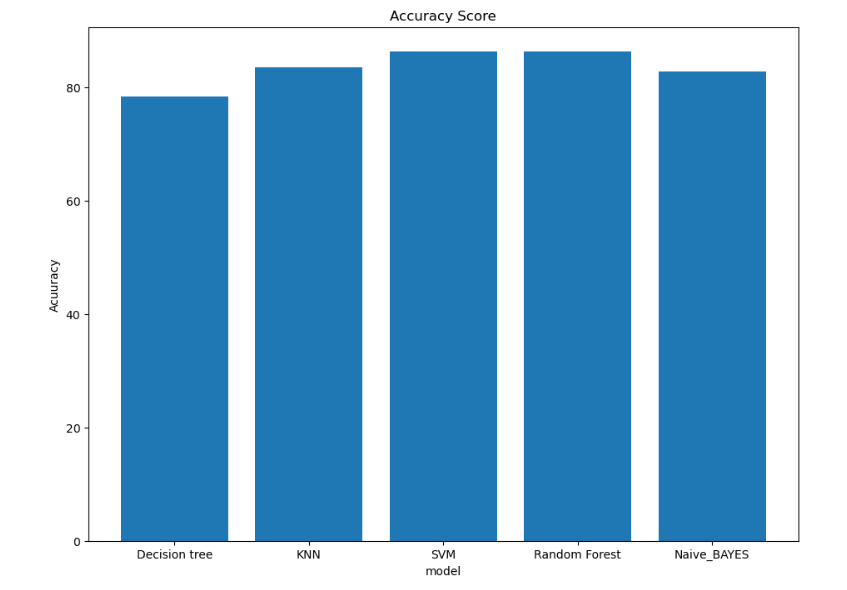
Ensure all sources used throughout your project are properly referenced in this section. A consistent citation style should be followed for all references, including books, articles, web sources, etc.

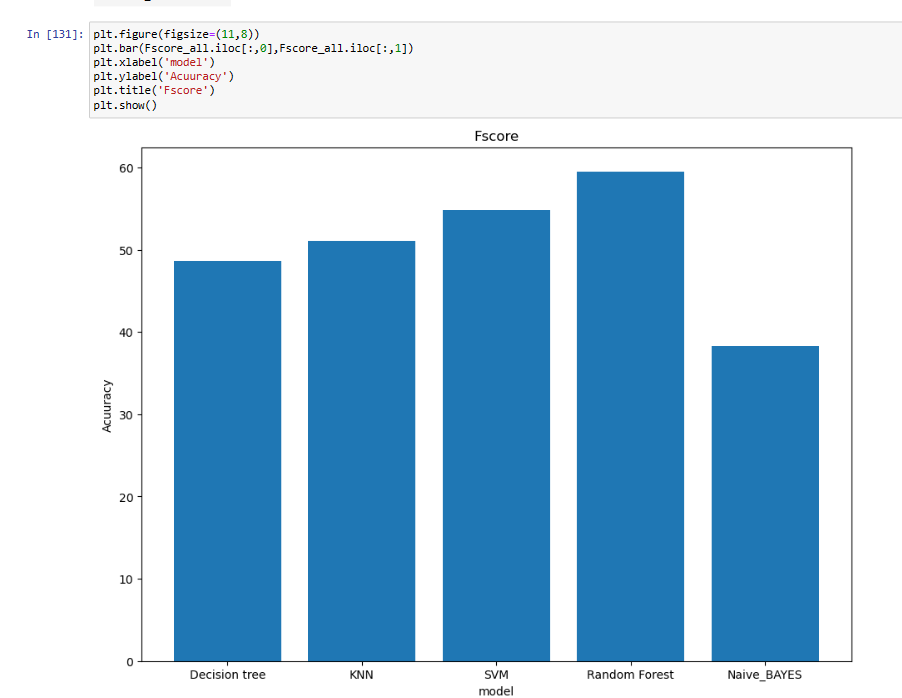
**Appendices**

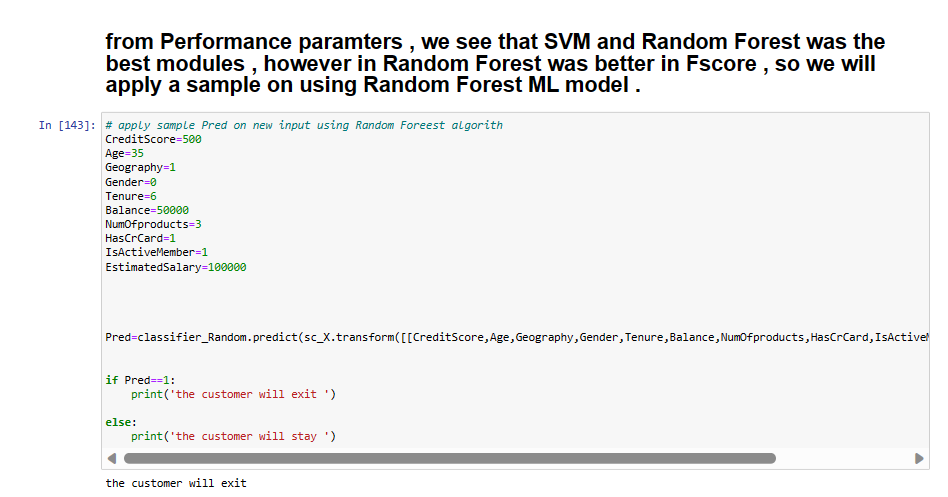
Include any additional materials that support your project but may not be necessary within the main text, such as raw data, code snippets, additional charts, interview transcripts, etc.

‘’ after completing the implementation of machine learning model and extract the confusion matrix then get the performance parameters , Accuracy Score and F1-Score .

We have found that SVM and Random Forest are the best modules to predict data .







**Deliverables:**

* A detailed project report documenting all the elements discussed above. **DONE**
* An executable code file along with any necessary files required for your program's execution. **DONE**
* A presentation that effectively summarizes your project and communicates your findings and insights to a non-technical audience. **DONE**
* Add your project in to the given GitHub Repository

**Additional Information:**

* The project should be completed within 4 weeks.
* Students will gather their own datasets relevant to the project.
* The project can be completed using any data science tool, programming language, or software.
* Students are encouraged to explore beyond the provided guidelines and investigate other potential aspects of data analysis.