**PROJECT REPORT: MY\_ALGEBRA\_PROJECT**

**Introduction:**

The "My Algebra Project" applies machine learning techniques to analyze and classify music data. The project focuses on preprocessing a dataset, reducing dimensionality using Principal Component Analysis (PCA), and training multiple classification models.

**Objectives :**

* Clean and preprocess the dataset by handling missing values and encoding categorical features.
* Apply feature scaling and dimensionality reduction for better model performance.
* Train and evaluate multiple machine learning models for genre classification.
* Visualize data relationships and model results.

**Methodology**

**1.Data Loading and Preprocessing:**

* The dataset (music\_dataset\_mod.csv) is loaded and examined for missing values.
* Missing values in the Genre column are removed.
* Features are standardized using StandardScaler.

**2.Exploratory Data Analysis (EDA):**

* A correlation heatmap is generated to understand feature relationships.
* PCA is applied to retain 85% of the data variance.

**3.Model Training and Evaluation:**

* The dataset is split into training and testing sets (70/30 split).
* Two models are trained: Logistic Regression and Support Vector Machine (SVM)
* Model accuracy and classification reports are generated for performance assessment.

**4.Genre Prediction for Missing Values:**

* Missing genres are predicted using the trained Random Forest model.

**Results**

* PCA reduced the number of features while preserving significant variance.
* The models achieved varying levels of accuracy, with SVM and Random Forest performing best.
* Missing genre values were successfully predicted.

**Conclusion**

This project demonstrates the effectiveness of machine learning in analyzing and classifying music data. The use of PCA for dimensionality reduction improved efficiency, and model evaluation helped identify the best-performing classifier.

**Contact**

For any questions or contributions, reach out to moivre20041@yahoo.fr.