**MACHINE LEARNING**

**FINAL PROJECT**

**INTRODUCTION:**

For the motive to resolve complicated issues that have been once not possible to clear up with traditional strategies, gadget gaining knowledge of has grown in significance. Our purpose is to use [dataset] to clear up the prediction problem for [target variable] on this challenge. [Give background information about why this problem is significant—e.g., predicting health outcomes, increasing customer retention, etc.] is the driving force at the back of the project. Our objective is to improve forecast accuracy and efficiency thru the utilization of machine learning techniques, which can yield enormous insights for [particular domain].Using a number of supervised getting to know tactics, together with [discuss any particular algorithms like Logistic Regression, Linear Regression, Decision Trees], the mission's major goal is to version the relationship between [predictor variables, such age, height, etc.] and [target variable].The primary purpose is to create a model that may appropriately forecast the future and generalize nicely to information that has no longer but been located.

**Methodology:**

**Dataset Description:**

The dataset applied for this work become provided by using the [source or file], and it contained [number of entries and number of characteristics]. The following are some of the substantial variables in the dataset:

A character's better body, expressed in centimeters, is the non-stop variable of top.  
Age: A continuous variable that represents an man or woman's years of age.  
Exercise Level: A categorical variable with values starting from Low to High that indicates how frequently someone engages in bodily activity.  
Weight: A continuous variable that expresses someone's weight in pounds, weight serves as the aim variable for prediction.

**Preprocessing Steps:**

The data went via a number of preparation tiers earlier than being carried out to any machine studying models:  
  
Managing Missing Data: Depending on their significance, rows containing lacking information were both removed or imputed.  
Categorical Encoding: To permit its inclusion inside the model, the specific variable Exercise Level became encoded the use of One-Hot Encoding.  
Scaling and Normalization: To assure consistency in importance, that's vital for some system mastering strategies, the continuous variables Height and Age were scaled the use of StandardAero.

**Machine Learning Algorithms:**

The statistics went via some of practise levels earlier than being carried out to any system learning fashions:  
  
Managing Missing Data: Depending on their importance, rows containing lacking facts were both eliminated or imputed.  
Categorical Encoding: To allow its inclusion in the version, the specific variable Exercise Level was encoded the usage of One-Hot Encoding.  
Scaling and Normalization: To guarantee consistency in significance, which is important for some system mastering techniques, the non-stop variables Height and Age were scaled the use of Standard Scaler.

**Evaluation Techniques:**

The following measures were used to evaluate the fashions: The average of the squared discrepancies among the expected and real values is called the imply squared blunders, or MSE. The percentage of the established variable's variant that may be anticipated from the independent variables is expressed as the R-squared statistical metric.

**Discussion:**

The effects reveal that the [best-performing model] supplied the very best accuracy in predicting the load primarily based on peak, age, and exercise stage. However, there are several factors that might be progressed:  
  
Model Limitations: [Discuss any limitations like overfitting, insufficient data, etc.].  
Feature Importance: Some features, consisting of [activity level], are thought to be more crucial in predicting weight because of their better affiliation with the objective variable.  
Refinements: Future studies need to become aware of additional elements in addition to diet, life-style, or genetics, as this can also improve the prediction model.

**CONCLUSION:**

This mission effectively developed a gadget studying model that can forecast weight relying on some of critical variables, including age, height, and diploma of workout. Even though the effects have been encouraging, in particular with [best model], there is nonetheless room for improvement, especially with regard to function choice and version complexity. The version might get even more correct with a larger and more various dataset, which might make it useful in situations like health tracking in the actual global.

**References:**

 Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer – A comprehensive resource on machine learning techniques, including regression and classification models.

 Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer – A classic reference on predictive modeling and machine learning.

 Pedregosa, F., et al. (2011). "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, 12(Oct), 2825-2830. – Describes the machine learning library used for model implementation.

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated