**Final Portfolio Piece (Project) Idea – Body Performance Data**

**1. What are you doing? What is the question?**

People have a sedentary lifestyle and have less physical exercise than earlier in today's world. We want to classify the person's body health with parameters such as age, gender, weight, vitals sign information, and the amount of exercise they perform. We will calculate their body class based on A to D. A being the best and D the worst.

**2. Data? Where is the data?**

This is data that offers measurement information for each item of the National Sports Promotion Corporation's national fitness measuring data for the Seoul Olympic commemoration. The data columns are age, gender, height, BMI, sit up performed by the individual

<https://www.kaggle.com/kukuroo3/body-performance-data>

<https://www.bigdata-culture.kr/bigdata/user/data_market/detail.do?id=ace0aea7-5eee-48b9-b616-637365d665c1>

**3.Which Algorithms? What approaches?**

1. Logistic Regression Training Accuracy: LogisticRegression(random\_state=101)
2. Decision Tree Training Accuracy: DecisionTreeClassifier(criterion='entropy', random\_state=101)
3. SVC Training Accuracy: SVC(random\_state=101)
4. NN Training Accuracy: KNeighborsClassifier(n\_neighbors=3)
5. GradiendBoosting Training Accuracy:
6. Random Forest Training Accuracy: RandomForestClassifier(criterion='entropy', n\_estimators=10, random\_state=101)
7. GradientBoostingClassifier(random\_state=101)
8. AdaBoosting Training Accuracy: AdaBoostClassifier(random\_state=101)
9. XGBoosting Training Accuracy: XGBClassifier(base\_score=0.5, booster='gbtree', class\_weight=None

**4.Evaluation? How do you know it worked?**

We will be evaluating by following the below steps:

1. Data Preparation
2. Model Training
3. Model Optimization
4. Model Evaluation:
   1. Classification Report
   2. Precision
   3. Accuracy
   4. ROC Curve
   5. Cross Validation on the dataset
   6. Tuning the Hyper parameters

After all of the machine learning stages have been finished and all of the algorithms and models have been run, we will compare the accuracy and rank the algorithms based on their accuracy.