The development of code for applying Gaussian Naive Bayes (GNB) classifier for activity recognition from sensor data was not without its fair share of challenges which moved from how data was to be preprocessed to how the models were to be evaluated. Everything was addressed by passing through the subsequent stages and the program was written.

1. Data Loading and Exploration:

The dataset is a CSV file that contains sensor data from a Run or Walk activity, and is assumed to have this filename ‘run\_or\_walk.csv’. After using Pandas to load the data, we review the header by using data.head() to examine the first records of our first data. The dataset contains accelerometer and gyroscope readings where the activity label serves as the target variable.

1. Feature Selection and Target Variable:

As mentioned before, the target variable, activity, informs that which activity was performed, either run or walk. All other columns (sensor Readings) are treated as features. The dataset is treated such that the independent variables are sensor readings and the dependent variable is activities.

1. Model Training and Evaluation (Full Feature Model):
2. Our first step consists of training over all sensor features, that is, both accelerometer and gyroscope data. Using the module train\_test\_split, the dataset was split into 70% for training and 30 % for testing. The GNB classifier was then fitted to the training set and the model was evaluated using accuracy, precision, recall and F1 scores on the test set.

4. Model with only Accelerometer:   
After that we fitted a model that only employs the accelerometer data. The same training and evaluation steps are repeated.  
  
 5. Model with only Gyroscopes:  
In the same manner, a model that applies gyroscope data alone is trained and the results evaluated.  
  
 6. Why So? Evaluation and Model Comparison:  
Each model (full feature model, accelerometer-only and gyroscope-only) is presented together with the computed accuracy scores and their respective classification reports. This serves as an indication of the contribution of each of the sensor types when distinguishing a “run” from a “walk”.  
  
Model Evaluation Metrics:  
• Accuracy: Number of correct predictions made.  
• Precision: The number of true positive predictions divided by the total number of positives predicted.  
• Recall: The number of true positive predictions divided by the number of true positive cases.  
• F1-Score: The harmonic average of precision and recall.  
• Support: The count of each class in the test data.  
  
Conclusion:  
  
The model that uses all parameters, that is all sensors data (the accelerometer and the gyroscope), is expected to give better results since it incorporates the information that two sensors have. The accuracy of the models that will be built using the accelerometer or the gyroscope alone, is expected to be lower since one sensor would not be enough to provide enough information to discriminate between the two activities. This approach builds a good basis for activity recognition which further can be improved with features engineering or more.