Flow Chart for Malware analysis work
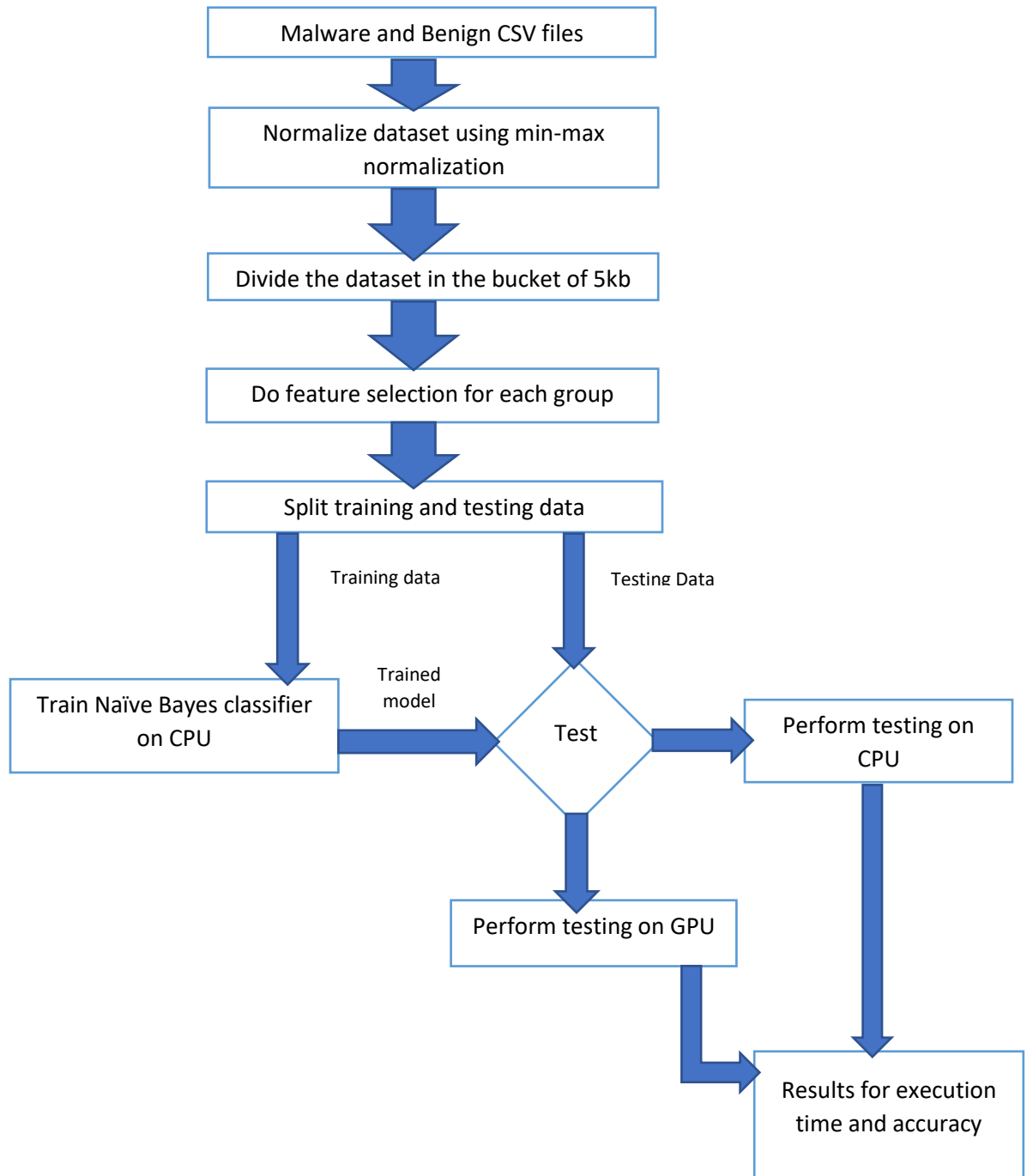
Naïve Bayes Classifier

Malware and Benign CSV files

⬇

Normalize dataset using min-max normalization

⬇

Divide the dataset in the bucket of 5kb

⬇

Do feature selection for each group

⬇

Split training and testing data

Training data   Testing Data

Train Naïve Bayes classifier on CPU → Trained model → Test → Perform testing on CPU

Perform testing on GPU

Results for execution time and accuracy

Neural Network

Malware and Benign CSV files

↓

Remove unwanted attributes

↓

Min-max Normalization

↓

Split Data into Training, testing and cross-validation data set (57-33-10)

Three Layer Neural Network

Input layer, 1809 input features

Hidden layer, 1809 PReLU units

Output Layer, I sigmoidal unit

Cross-validation data

Testing data

10-Fold Cross-validation

↓

Training accuracy calculation

Trained Model

Testing

Testing on CPU

Testing on GPU

Final score calculation for confusion matrix, serial and parallel execution time