A Framework for Sensor Based Hand Gesture Recognition Using Machine Learning

Exam Roll: **16**
Reg no. : **2015-516-695**
Session: **2015-16**

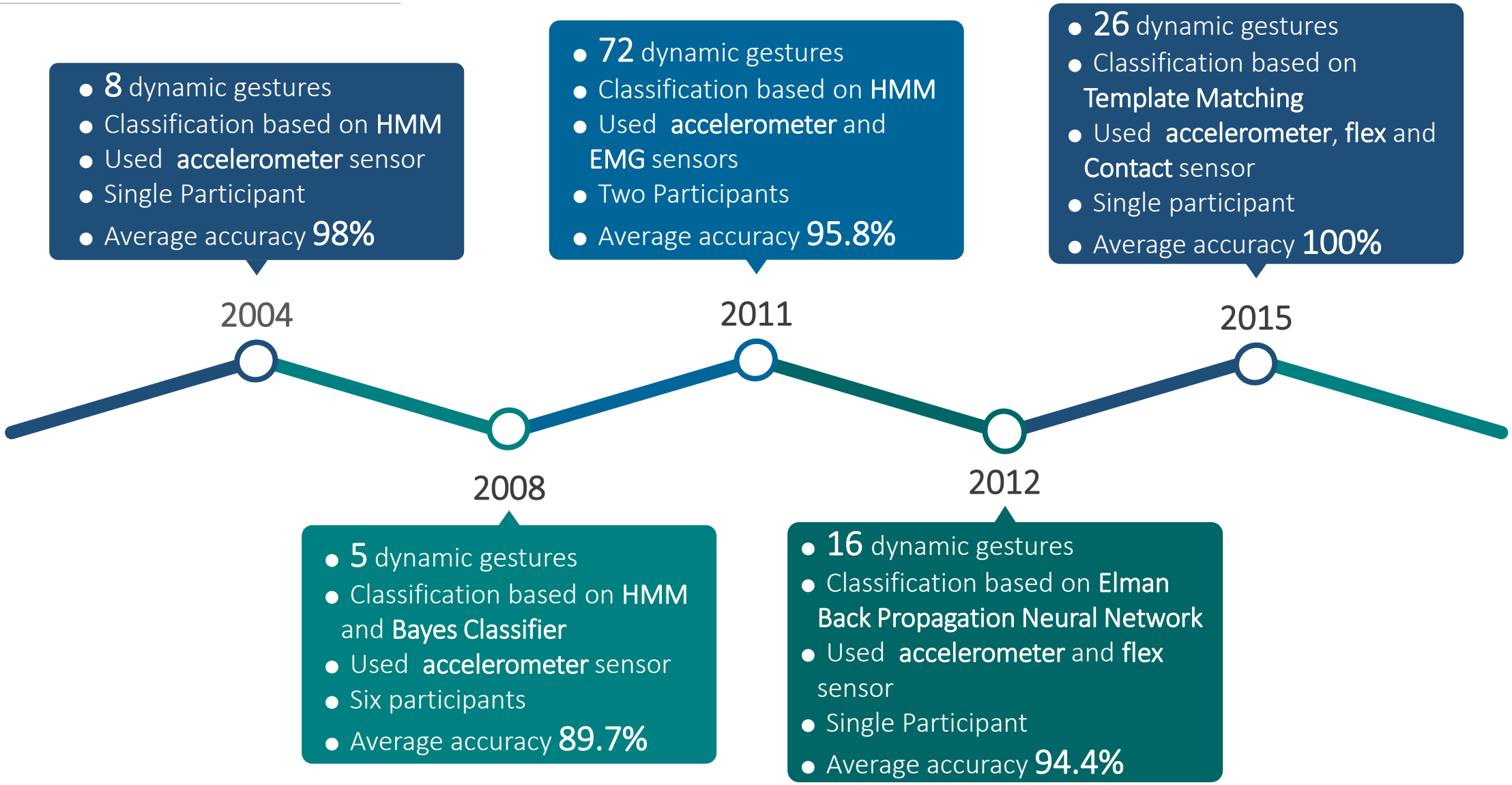
Exam Roll: **37**
Reg no. : **2015-216-698**
Session: **2015-16**

Department of Electrical and Electronic Engineering
University of Dhaka

Goal

- Static Hand Gesture Recognition Device
  - Dataglove
  - Desktop Software
- Making a Dataset
  - Diversity
  - Statistically Applicable Dataset Size
- Comparison of Accuracy of Four Classification Models

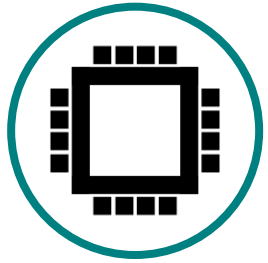
Related Works



MCU and Sensor Selection

3 Sensors

- 5 Flex Sensors
- 3-axis Accelerometer
- 3-axis Gyroscope



ESP32

- Xtensa dual-core **32-bit** LX6 MPU
- **240 MHz** clock
- **4 MB** Flash and **520 KB** RAM
- **12-bit** ADC
- **WiFi** and **Bluetooth** connectivity

Our Dataset

30
Participants

- **21** Male
- **9** Female
- **21-32** yrs age group

14
Hand Gestures

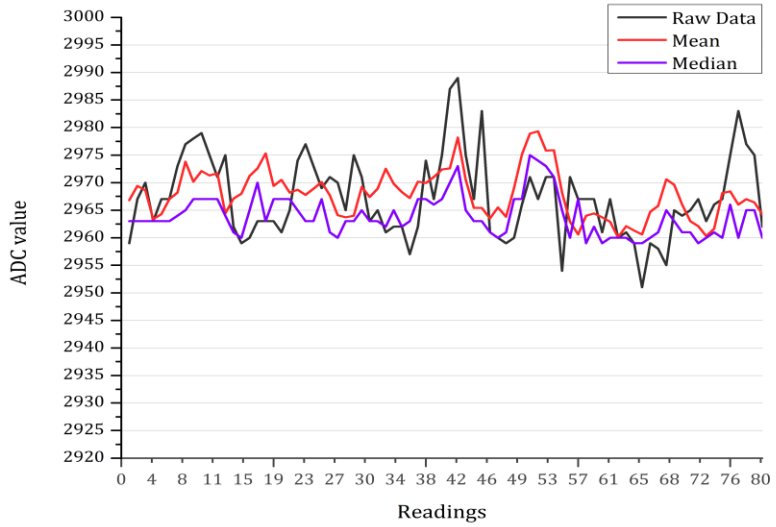
- **Static** gestures
- **10** instances recorded
- **8** features

Data Processing

Flex Sensor

Filtering & Normalization

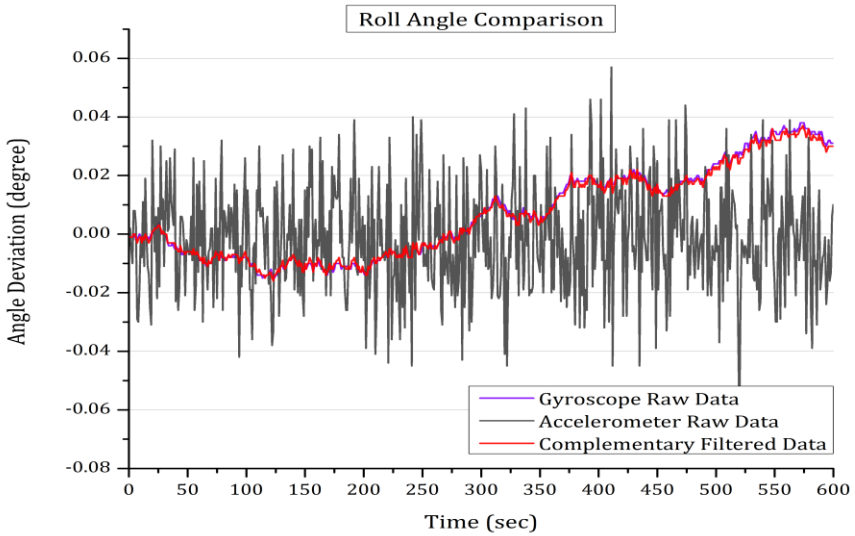
- Sampled every **10ms**
- **Median filter** is applied
- Values are **normalized**



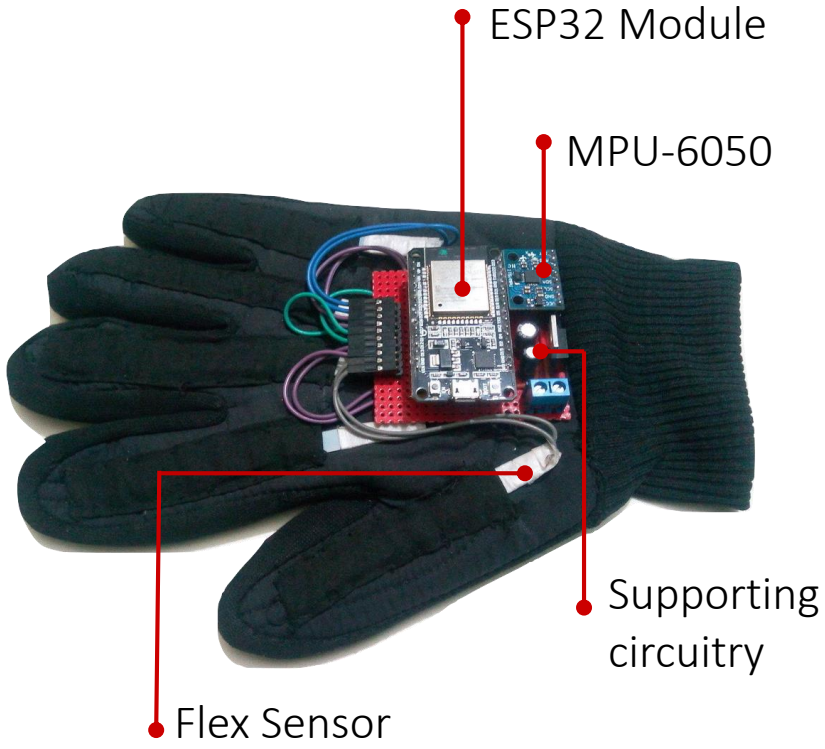
IMU Sensors

Filtering raw data

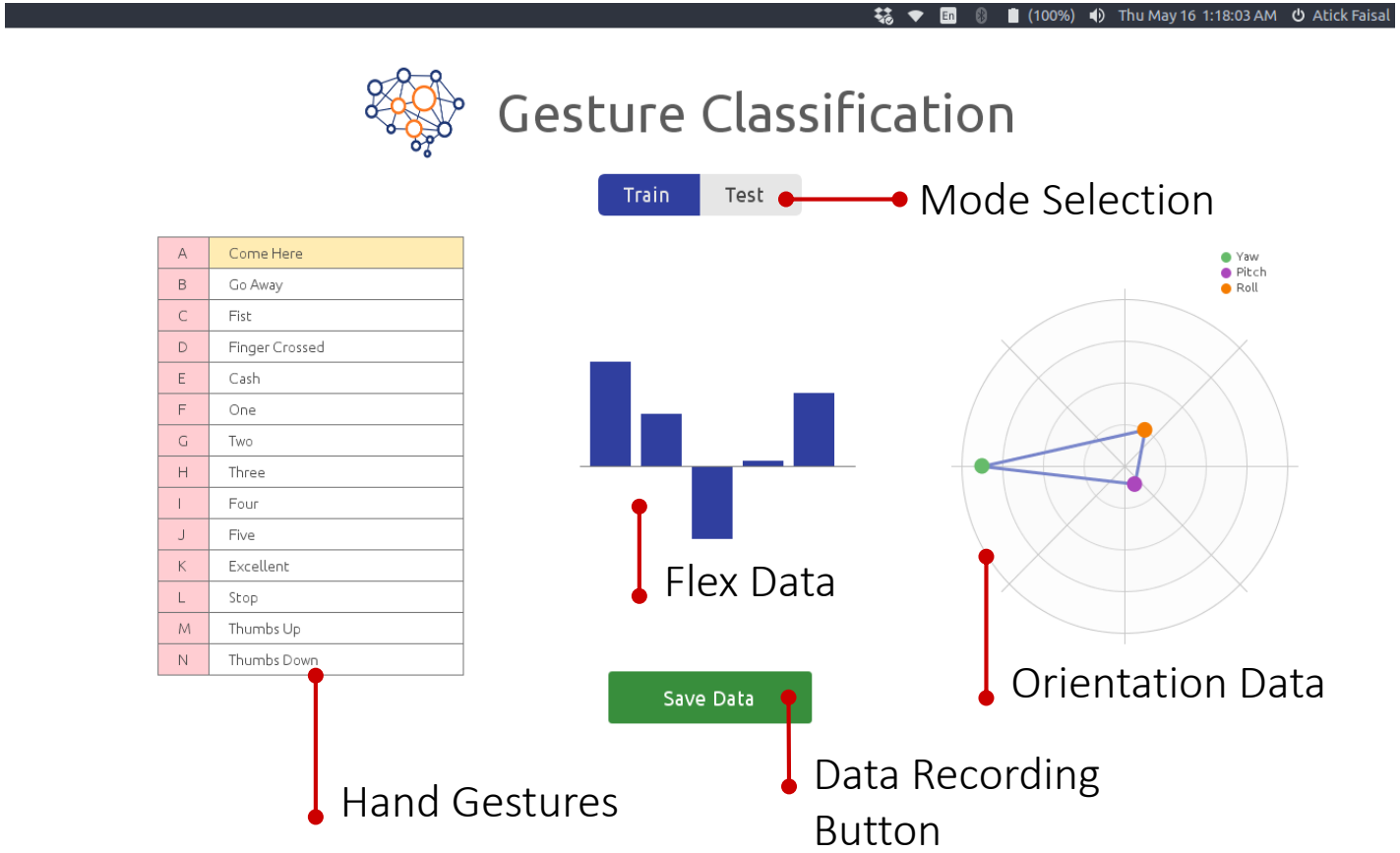
- Accelerometer readings are **noisy**
- Gyroscope values **drifts** over time
- Use of **Complementary filter**
- Remove yaw angle drift using **DMP**



The Device and the Software



Data Glove



Desktop UI

The Framework

Sensor Calibration

Flex sensors are calibrated and the values are normalized. IMU sensors are initialized.

Data Collection

Flex and IMU sensors readings are taken and sent to the desktop software. In the software data is verified and saved using the interface.

Data Preprocessing

Raw sensor data are preprocessed using different filtering techniques before training the model.

Training the Model

Different machine learning models are trained on the processed sensor dataset. Models are tuned by changing different parameters.

Making Prediction

The best fitted model is used to make the real time predictions.

Result Analysis

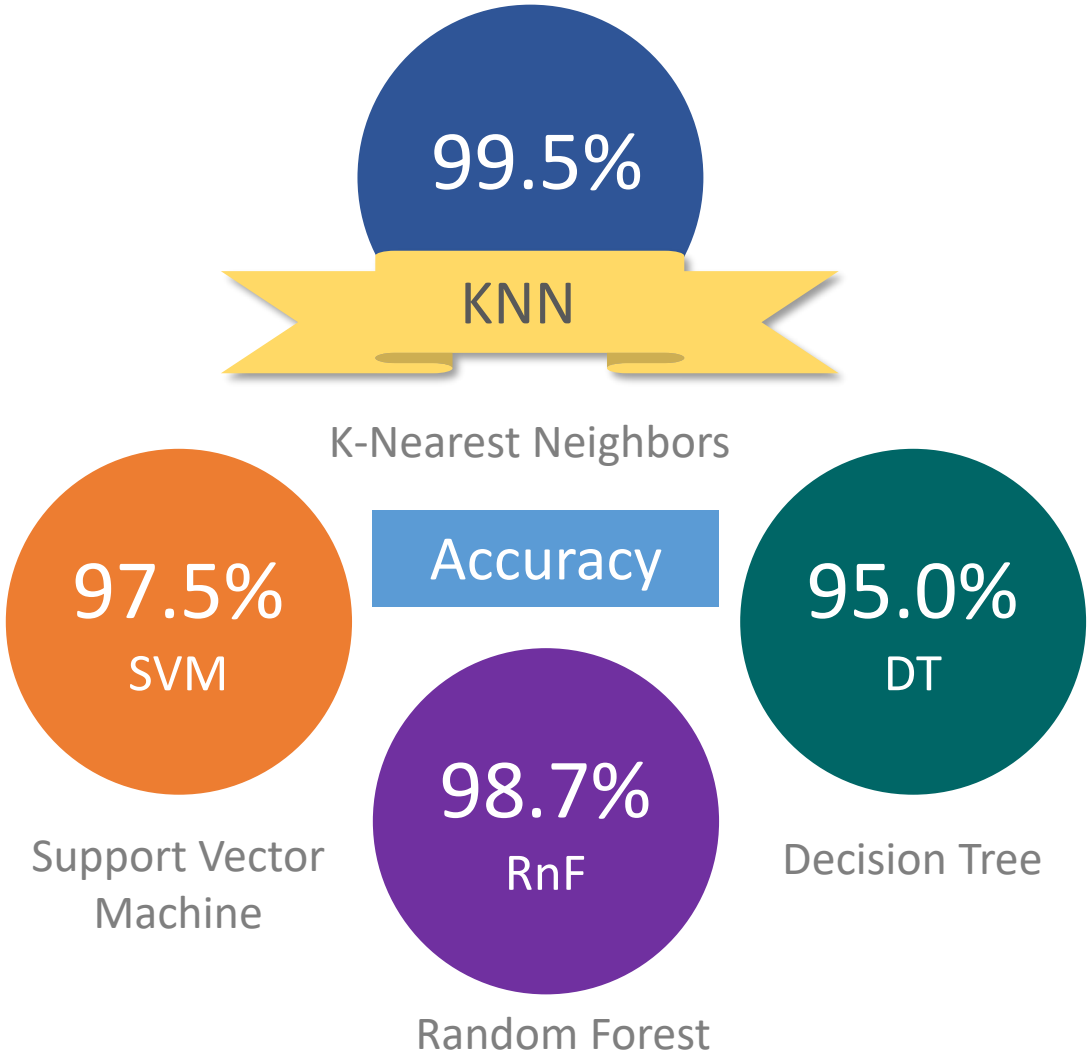
Evaluation Process

- Confusion Matrix
- Precision and Sensitivity

$$Precision = \frac{TP}{TP + FP}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

- Accuracy using **Cross Validation**



Future Work

Extending the **Dataset**

Using **Kalman Filter** with IMU Data

On-device Classification

Bangla Sign Language Classification

Dynamic Data Classification

Thank You