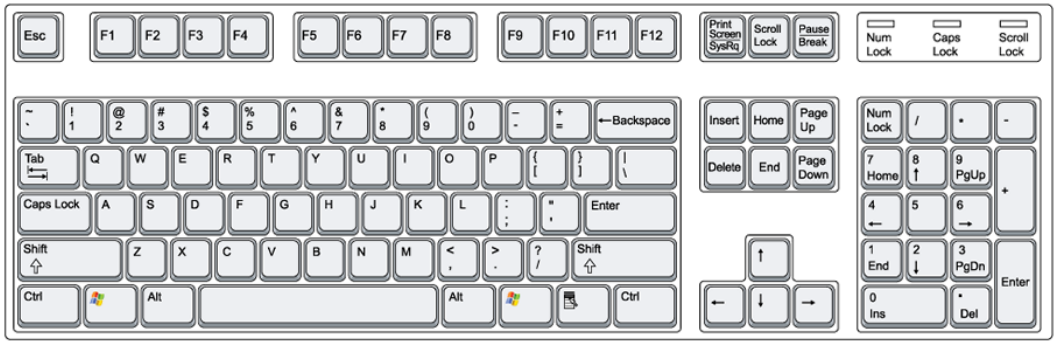


KEYBOARD ACOUSTICS

Huy Nguyen

Imagine this scenario…

This has already been done!

- Asonov, D., and Agrawal, R. “Keyboard Acoustic Emanations”. In *Proceedings of the IEEE Symposium on Security and Privacy* (2004), pp. 3–11.
- Distinguishing between 30 keys
  - 79% candidate with highest probability selection, 88% in the highest three
- Training on one keyboard, test on another
  - 50-52% in the highest four candidates

Why does this work?

- Each key consists of:
  - A head (the part you see)
  - A piece of rubber
  - An intermediate plastic piece
  - An electrical switch that closes a circuit
- Different parts of the keyboard plate produce different sounds when a nearby key is pushed

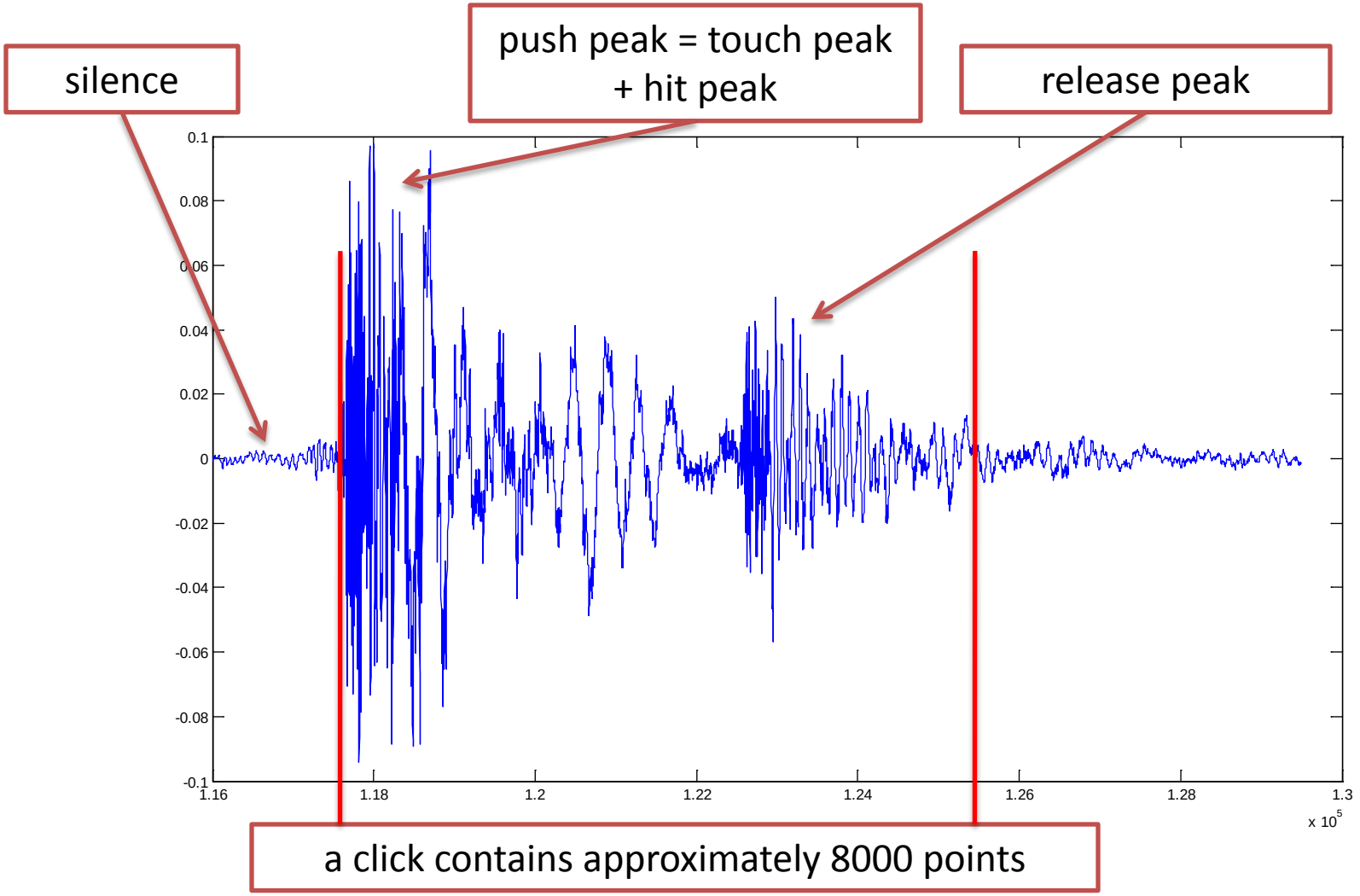
Our experiment

- Record clicks via Skype into .WAV file
- Extract individual keystroke from .WAV file
- Extract features from keystroke's audio signal
- Construct Neural Net to classify different keys based on extracted features
- All source code are written in Matlab
- Use Matlab Neural Network Toolbox

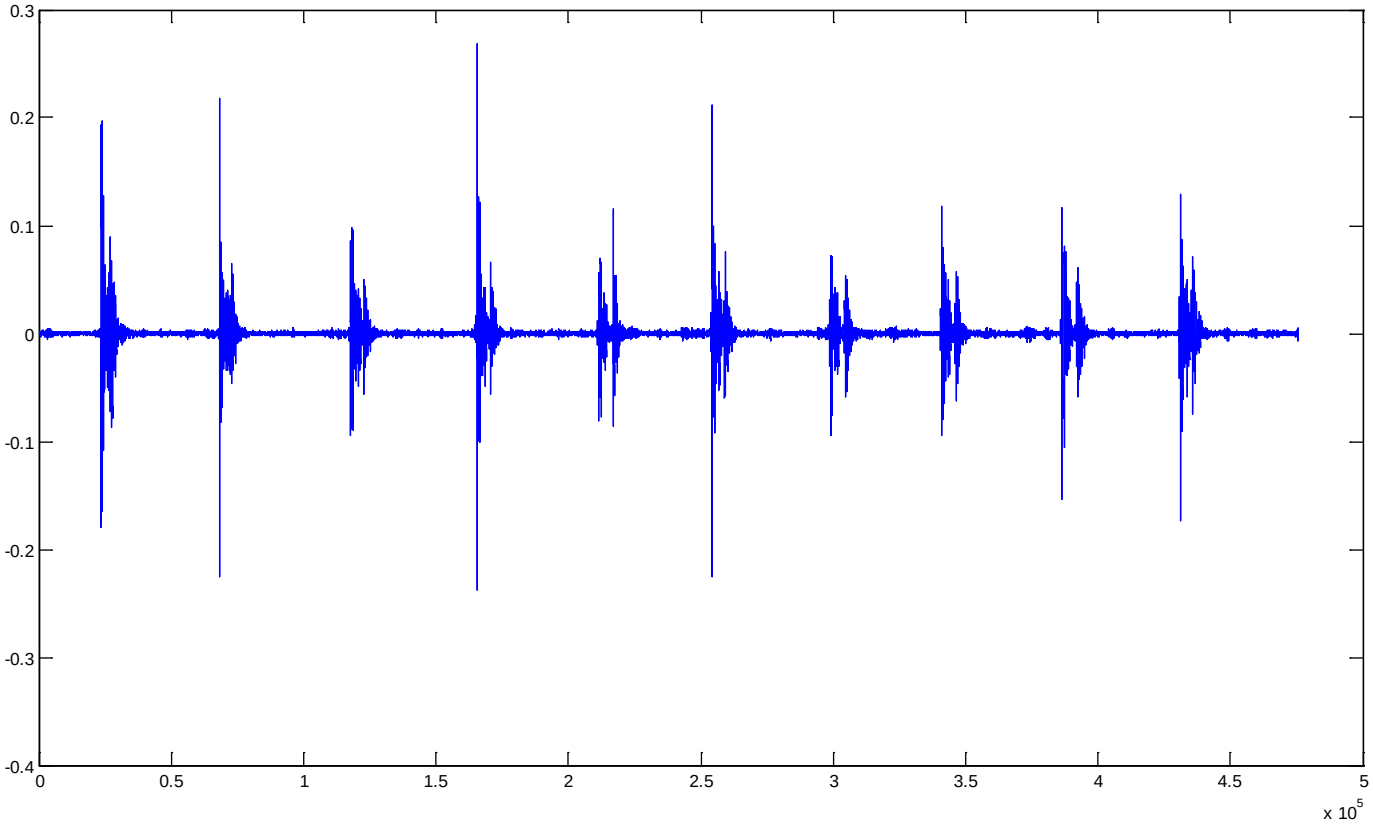
Detailed .WAV file information

- Each file is a recording of multiple keystrokes of the same key
  - 20 or 100 keystrokes in a file
- Audio Format: PCM
- Channels: Mono (1 channel)
- Sampling Rate: 44100 samples/s
- Bits/sample: 32 bits

An individual keystroke

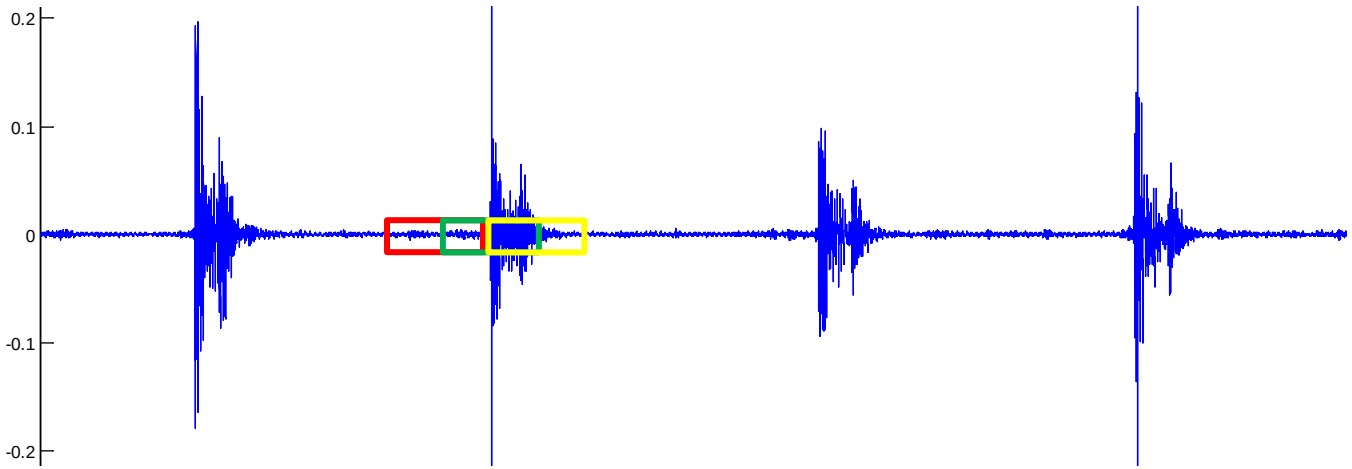


A .WAV file with 10 keystroke



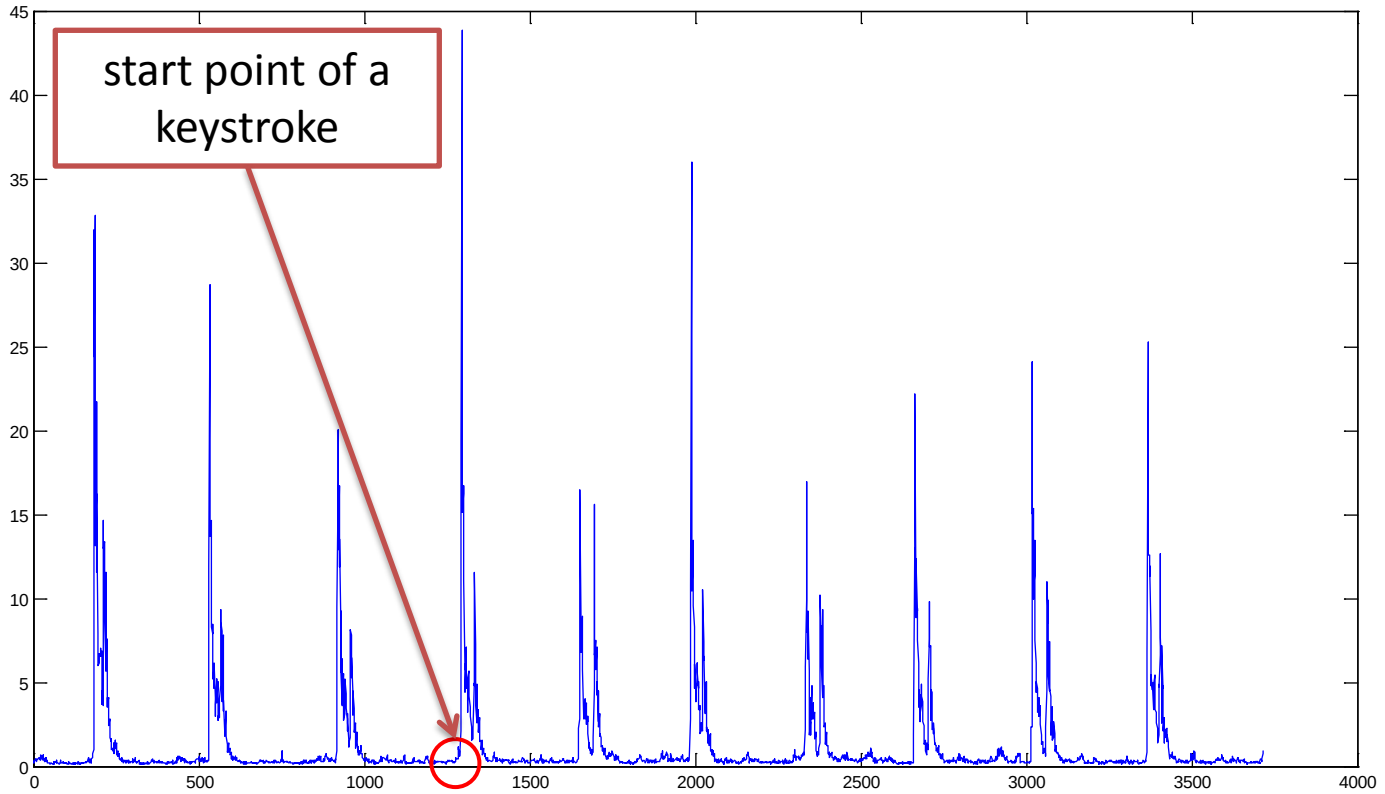
Keystroke extraction

- Data of the .WAV file is loaded into an array
- The array is sampled to discrete overlap windows
- Each window is put into a FFT procedure
- Sum of magnitude of computed FFT points is returned $\rightarrow$ energy level of the window



Energy levels for 10 keystrokes

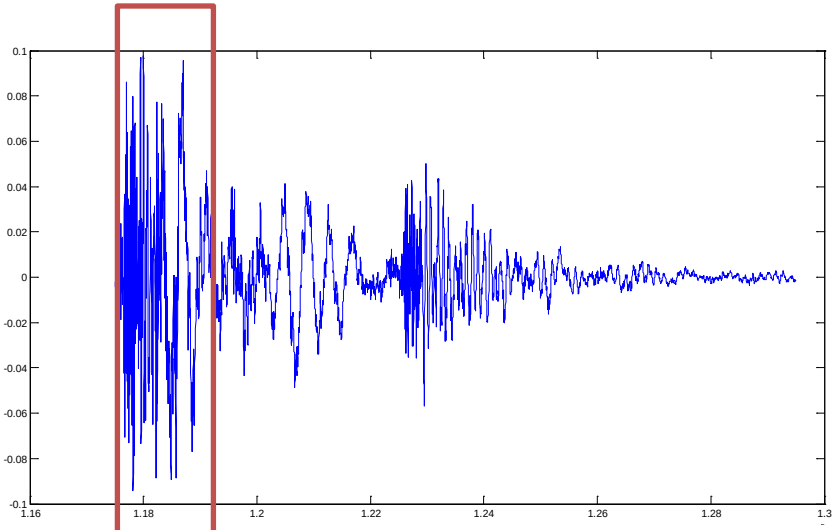
- Detect first point of keystroke's energy → convert into start point of keystroke's signal

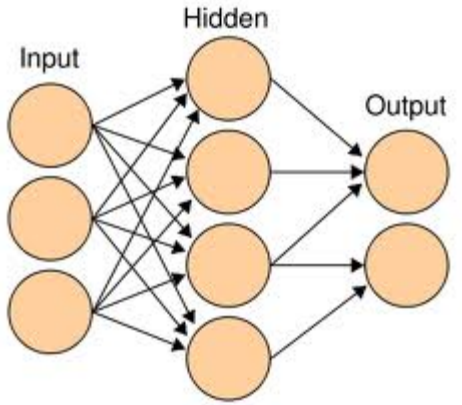


Features extraction

- Raw audio signal is not good input for classifier
- Features are FFT points
- Run FFT procedure over keystroke's signal
- Extract only FFT point in interested frequency range, i.e. 20÷4000Hz

Feature vector of a keystroke

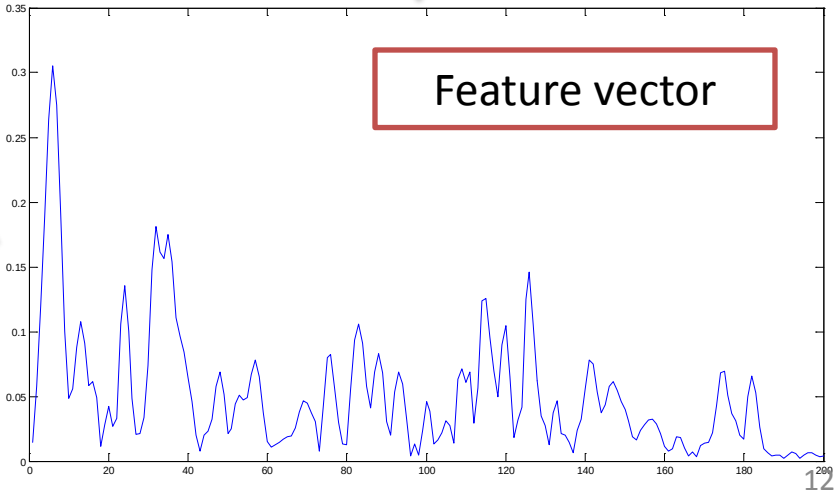




Input of Neural Net

extract features
for push peak only

Feature vector



Neural Network Data

- Inputs a vector of 200 FFT points
  - Range of [20, 4000]Hz with 20Hz unit
- Targets listed as vectors of 27 bits
  - Each vector has 1 one, 26 zeroes
  - 100000…. = A, 010000…. = B, ...000001 = _
- Hidden layer with 100 nodes
- Net trained on training data
  - Training data is then again split into 80% train, 10% validation, 10% test and trained
  - These subsets allow the network to determine when to stop

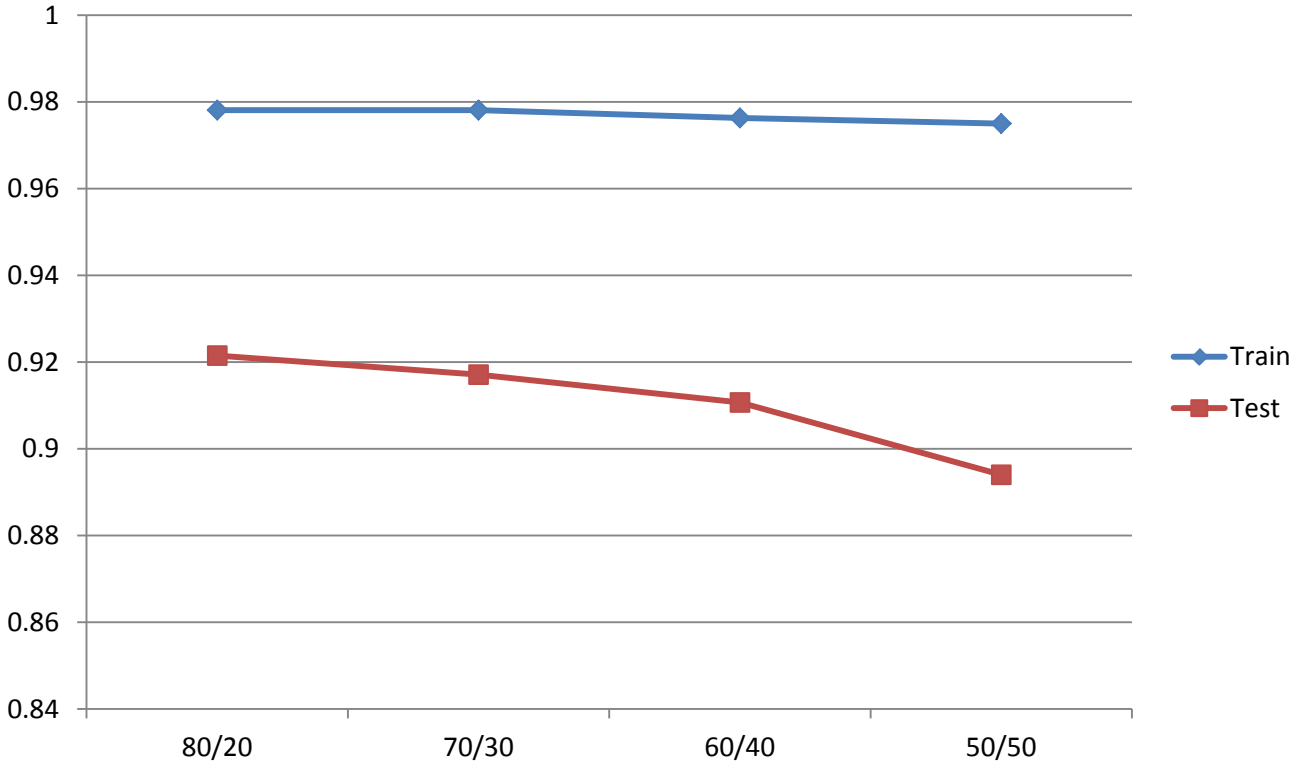
Data Collection

- Two computers set up with Skype
- Microphone suspended 1 foot above keyboard on carpeted floor to reduce noise
- Recorded with Audacity

Experiment with 1 keyboard

- For each key, 100 clicks are recorded into a .WAV file
- 27 keys, A-Z and space bar
- Changing train/test proportion
  - 80/20, 70/30, 60/40, 50/50
- 5 runs each setting, record only runs that all 27 keys are recognizable
- Performance measured by accuracy
  - only largest probability element

Experiment with 1 keyboard



Experiment with multi-session recording

- The same keyboard, but clicks are recorded in different days
  - 5 sessions: Monday to Friday of a week
  - A new recording in next week
  - Only 20 clicks/key in each session
- Test 1: Mix all 6 sessions, train/test proportion is 80/20
- Test 2: Train on first 5 sessions, test on last session

Experiment with multi-session recording

- Test 1:
  - Train performance: 0.73
  - Test performance: 0.58
- Test 2:
  - Train performance: 0.76
  - Test performance: 0.50
- Data variety affects greatly to performance
- Low performance is also due to bad recording

Experiment with two keyboards

- 2 keys A and P
  - Train on keyboard 1, clicks are recorded in 2 sessions
  - Test on keyboard 2 of the same model
  - Recognition rate is > 90%
- 27 keys
  - Train only on 1 session, 100 clicks each key
  - Recognition rate is < 10% !

Future Works

- Get more data + Collect data in multiple settings/sittings
  - Improve data variety and data quantity
- Try different features
  - ceptrum
- Try different classifier
  - SVM
- Keystroke extraction is a nontrivial task
  - Start point, click VS. noise, quick typing...

Thank you!