Drum Samples Classification

Final project for MAS 622J/1.126J

Eyal Shahar
MIT Media Lab

- **Objective :** Classify individual drum sample
- **Project focus :** Find best method and features set

Classes

- Bassdrum
- Snares
- Hi-Hat
- Cymbal
- Clap
- Tom

Methods

- Support Vector Machine
- K Nearest neighbors
- Neural Network

Features

- **Pitch** – Average maximum FFT bin frequency
- **Decay** – Time from peak to 50% amplitude
- **Brightness** – high frequency energy
- **Rolloff** – low frequency energy
- **Roughness** – based on the frequency ratio of each pair of sinusoids
- **Irregularity** – degree of variation of successive spectrum peaks
- **MFCC** – Mel frequency cepstrum coefficients

Training and testing dataset distributions

	Training		Testing	
	Samples	**Percentage**	**Samples**	**Percentage**
Bassdrum	337	33.2%	224	18.9%
Snare	259	25.5%	251	21.2%
Hi-hat	195	19.2%	403	34.0%
Cymbal	116	11.4%	68	5.7%
Clap	97	9.5%	52	4.4%
Tom	12	1.2%	187	15.8%
Overall	**1016**	**100%**	**1185**	**100%**

Matlab GUI

- Training and testing sets management
- Saving, loading, feature extraction
- Model management
- Feature space view
- Progress bars
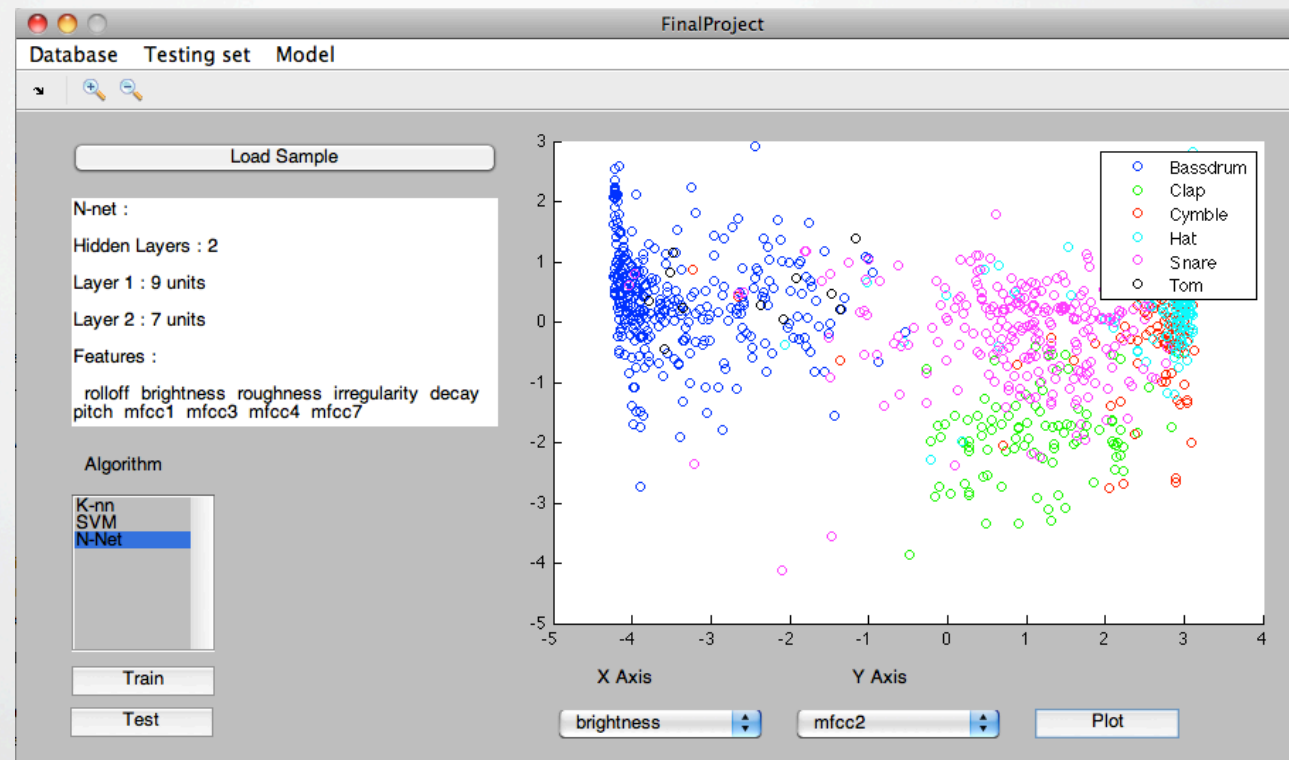
K Nearest Neighbors

- Tested $1 < k < 15$ with k taking odd values

- Leave-one-out validation

- Forward feature selection

Results : k = 9 using

- Brightness

- Irregularity

- Decay

- MFCC 1

- MFCC 2

- MFCC 3

- MFCC 5

■

Neural Net

- Used Matlab's neural net tools, including validation

- Performance estimation through MSE

- Tested 1 and 2 hidden layers

- 5:12 units per layer

Results

2 hidden layers with 9 and 7 units respectively

Selected features :

- Rolloff
- Brightness
- Roughness
- Irregularity
- Decay

- Pitch
- Mfcc 1
- Mfcc 3
- Mfcc 4
- Mfcc 7

■

Support Vector Machine

- Used Matlab built-in SVM tools

- Leave-one-out validation

- One SVM per class – One-Versus-All
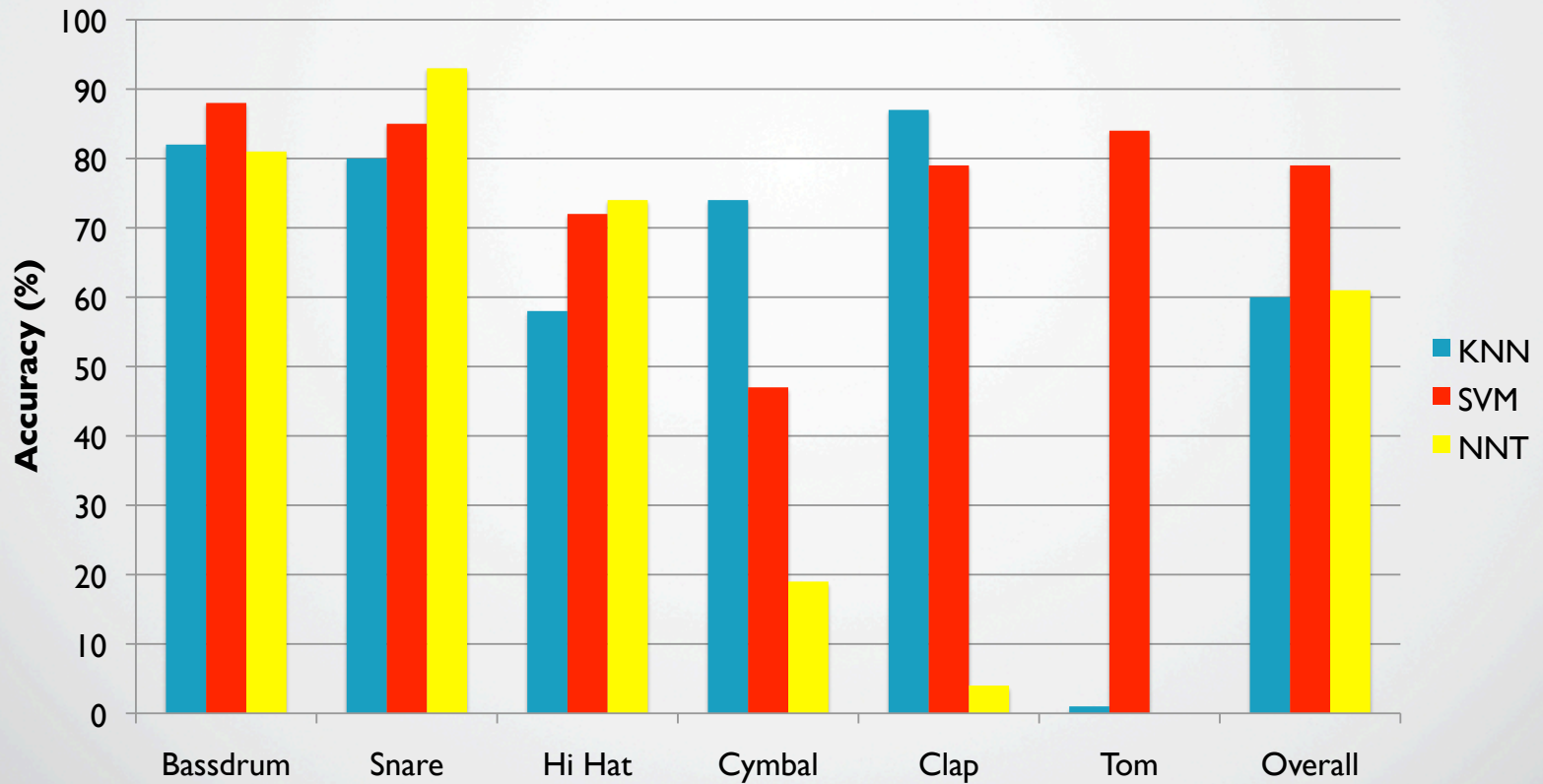
- Forward feature selection

Results

	Brightness	Rolloff	Irregularity	Roughness	Decay	Pitch	Mfcc1	Mfcc2	Mfcc5	Mfcc7	Mfcc11	Mfcc13
Bassdrum	✔						✔					
Snare	✔	✔				✔	✔		✔		✔	
Hi-hat					✔	✔					✔	
Cymbal	✔			✔	✔					✔		✔
Clap			✔					✔				
Tom		✔										

■

Performance

Method	Bassdrum	Snare	Hi-hat	Cymbal	Claps	Tom	Overall
KNN	82%	80%	58%	74%	87%	1%	60%
SVM	88%	85%	72%	47%	79%	84%	79%
N-net	81%	93%	74%	19%	4%	0%	61%

Improvements to be made

- Extend and manually validate training set

- Try temporal features approaches, namely HMM

- Implement multiple algorithms and majority selection

Future Work

- Subjective features extraction and sample retrieval

- Melodic instruments samples