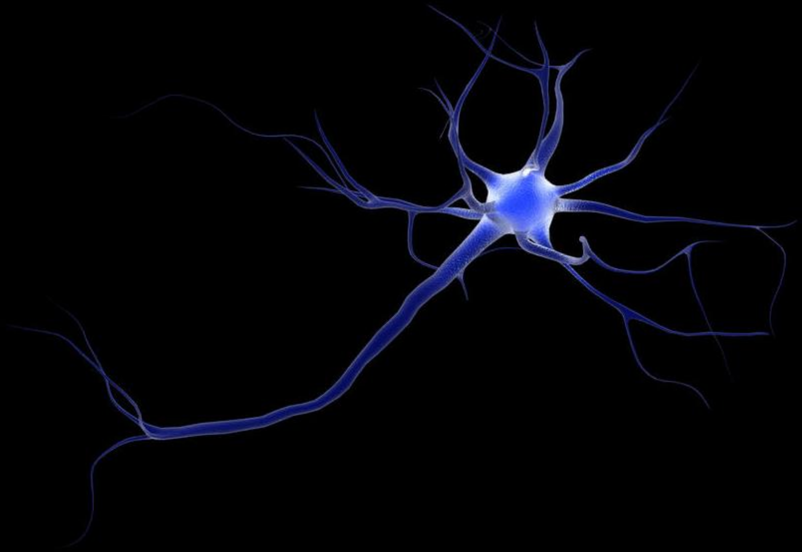
Neural spike sorting using deep learning based approaches

János Szalma, Tamás Nagy
Budapest University of Technology and Economics

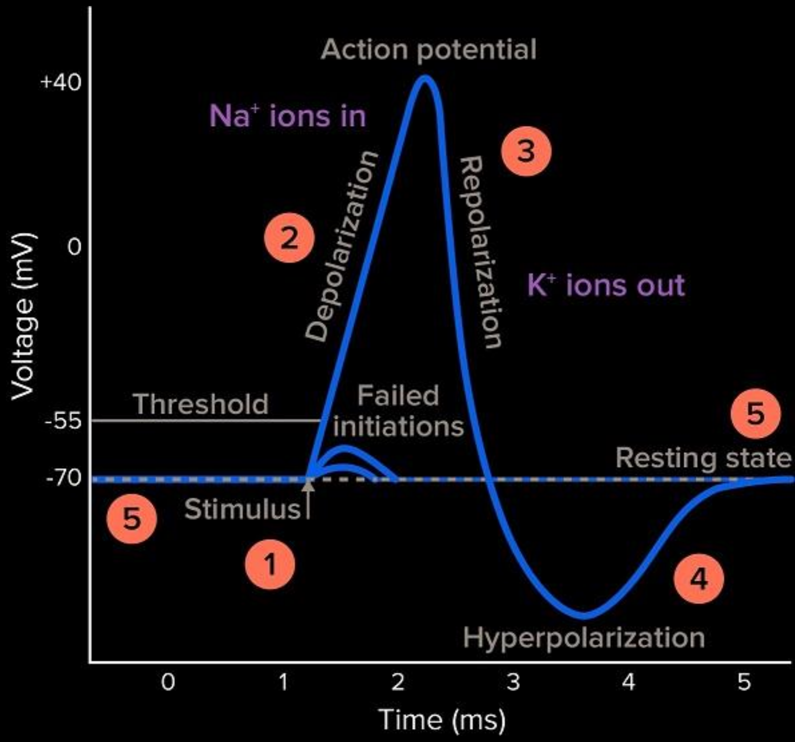


Introduction

Standard method to spike detection and sorting

Neural spikes

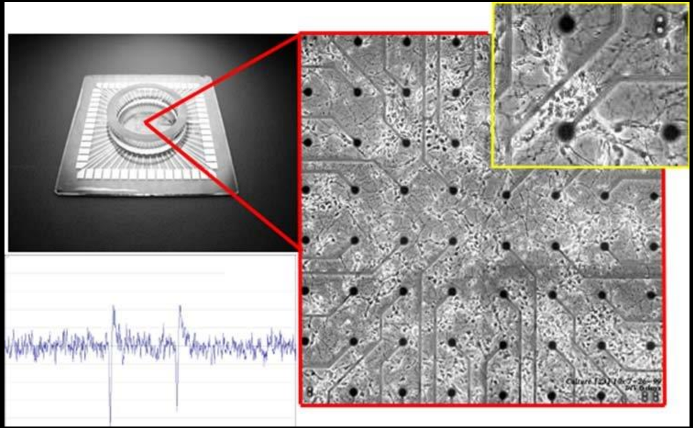
- during neural activity, individual neurons are firing
- inter-neuron communication is accomplished via neurotransmitters
- this needs voltage change to propagate along the axon
- we can measure the voltage change

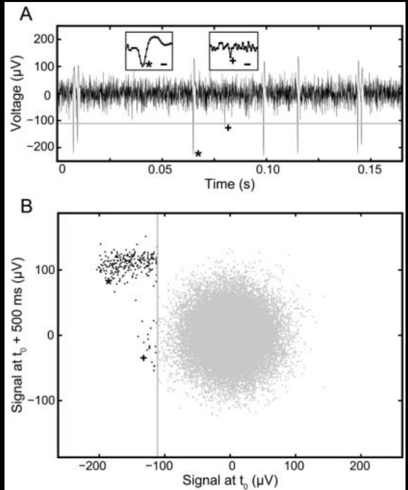


Spike detection

- the voltage is measured with external electrodes
- this method is prone to pick up environmental noise
- a single electrode can records the signals of multiple neurons

- detection: thresholding with the analysis of the event’s surroundings is usually sufficient

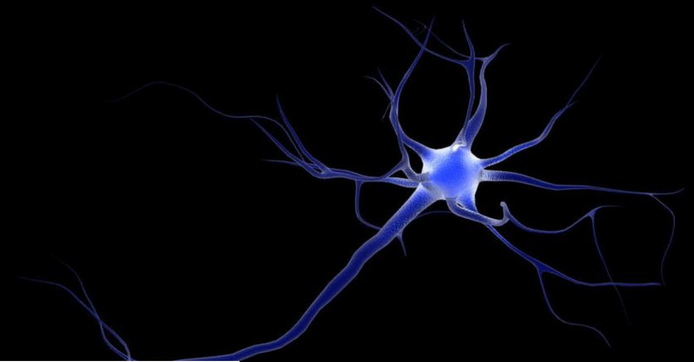


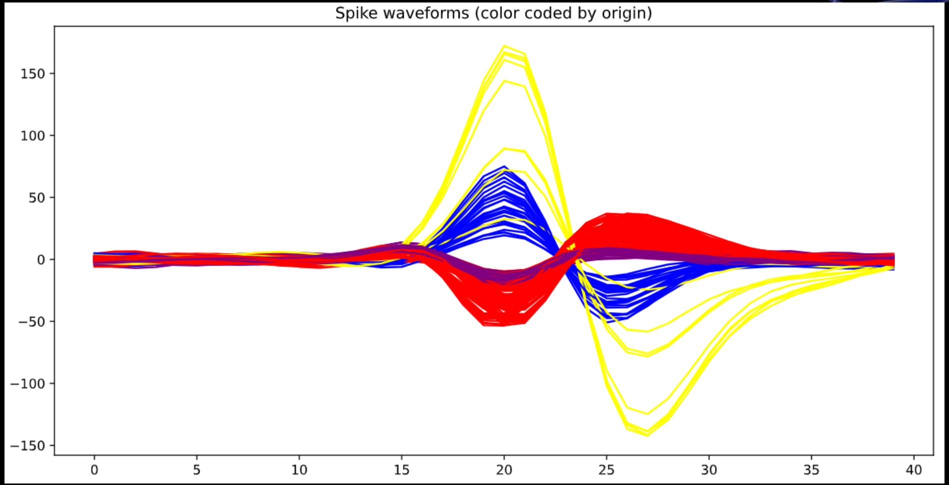


Data

Dataset

- artificial neural spikes
- 2.4 million spikes
- 40 time points per spike
- five different waveforms

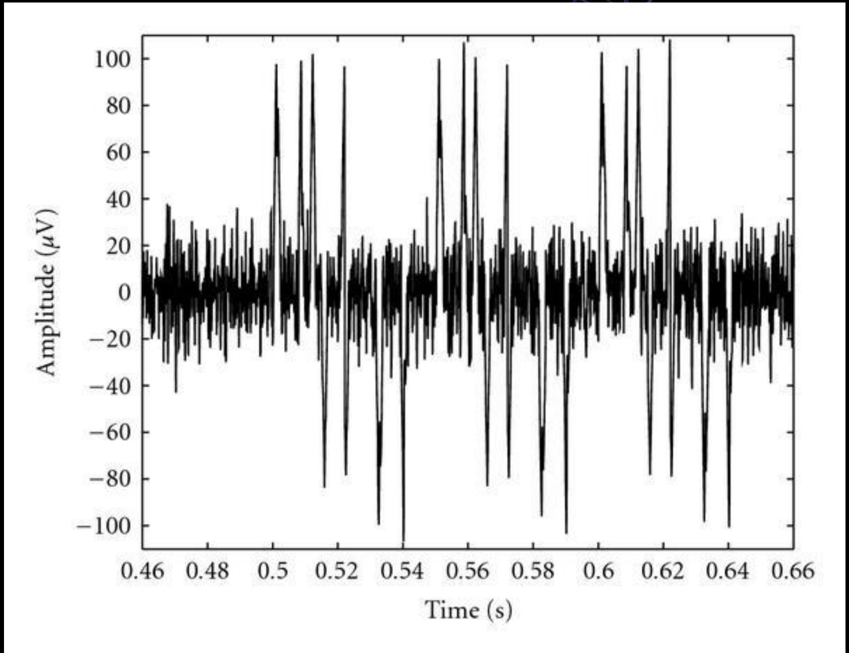




Preprocessing

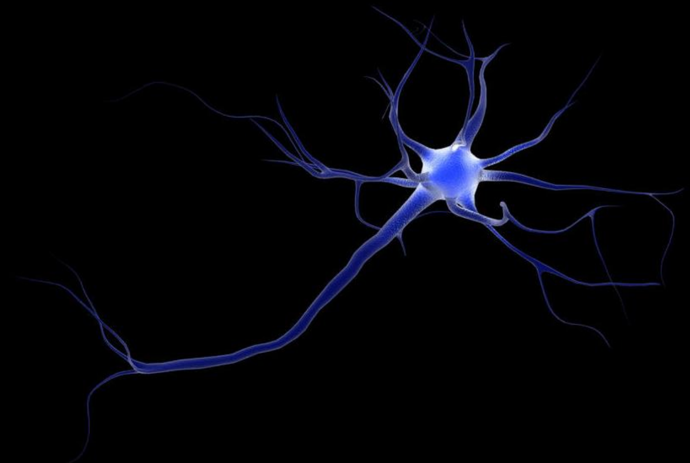
- shifting labels
- bandpass filter (150-2500 Hz)

- spike detection
  - AdaBandFlt noise estimation
  - threshold based detection
- preliminary results -> too easy?



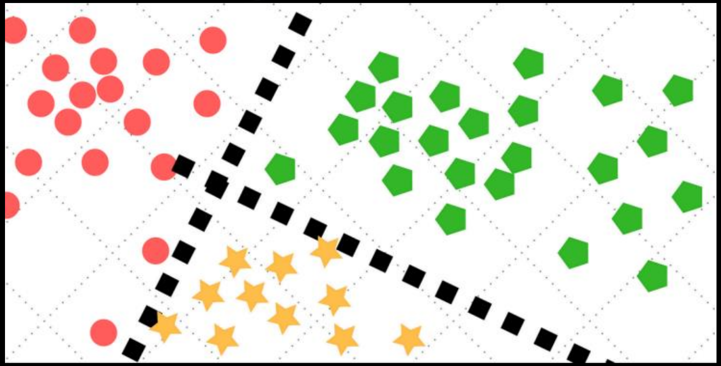
Classification

Classification with least data necessary



- 5 class classification
- validation: 5%
- test: 85%
- training: 0.1%-10%

- classifiers:
  - Logistic Regression
  - Random Forest
  - Convolutional Neural Network



Classification with least data necessary

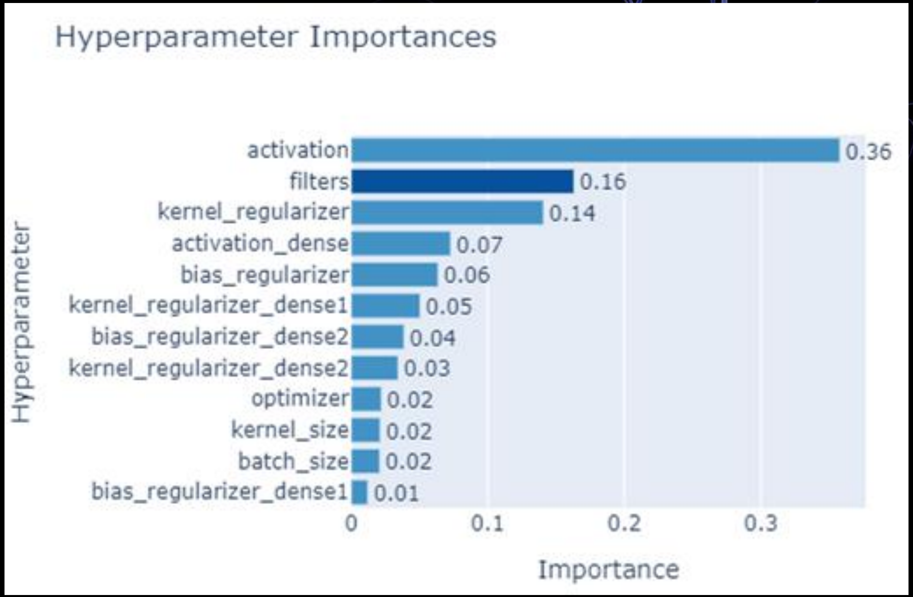
- focus on CNN optimization
- Tree Parzen Estimation
- stride: 1, padding: same
- batch size: [128,256]
- optimizer : [adam,nadam]
- metrics
  - balanced accuracy score
  - multiclass area under the roc curve

TABLE 1- CNN CLASSIFIER ARCHITECTURE

Hyperparameters	**Layers**		
	1D Convolutional	***Dense1***	***Dense2***
Filter	3-15	-	-
Kernel size	3,4,5	-	-
Kernel regularizer	None, l1, l1_l2	None, l1, l1_l2	None, l1, l1_l2
Bias regularizer	None, l1, l1_l2	None, l1, l1_l2	None, l1, l1_l2
Activation	Swish, sigmoid, relu	Swish, sigmoid, relu	Softmax
Output neurons	-	5	5

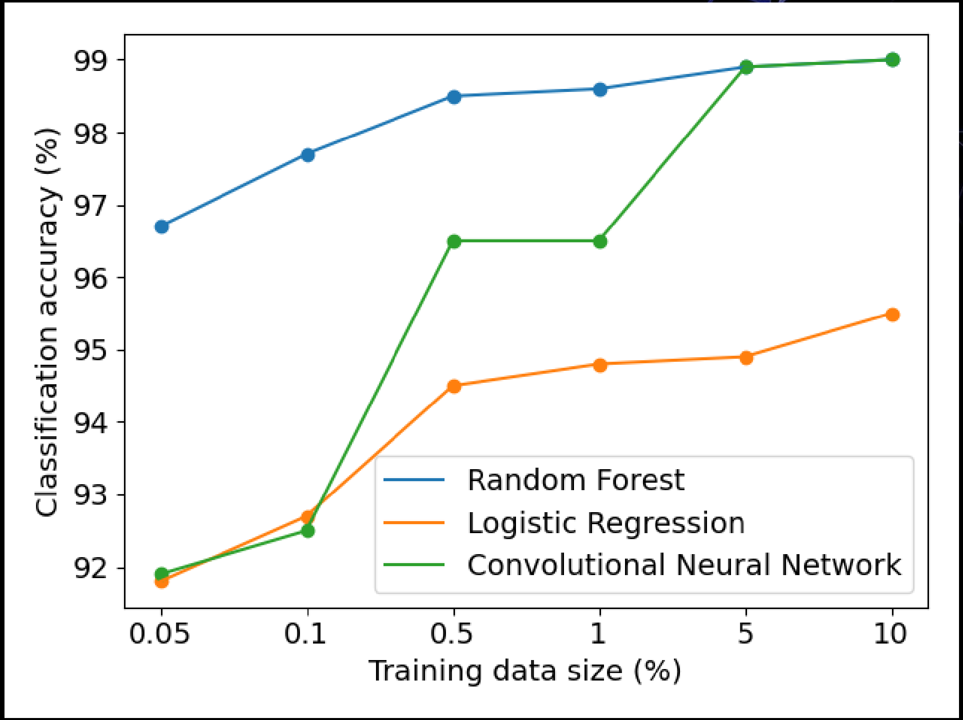
Hyperparameter optimization results

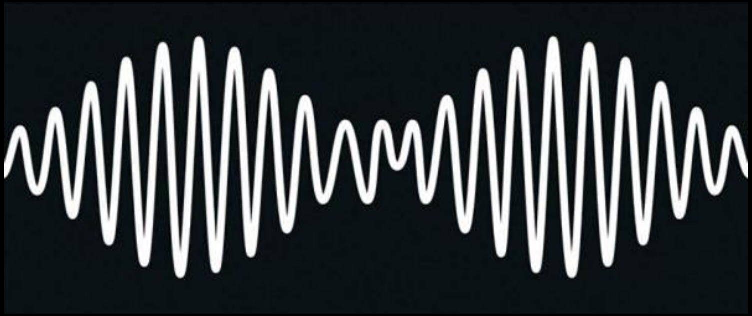
- Most parameters unimportant
- swish activation
- 13 filters
- no regularization



Classification results

- CNN needs at least 5% data
- LR shows worse performance
- RF best with small data size

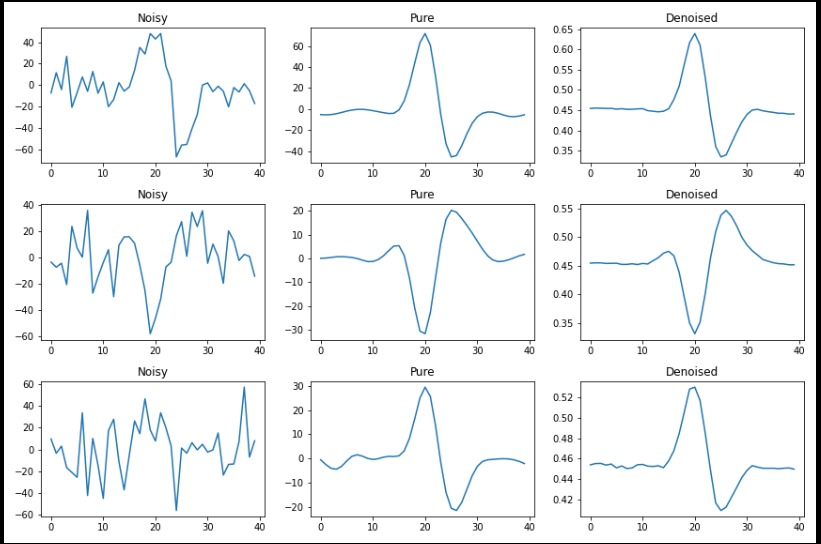


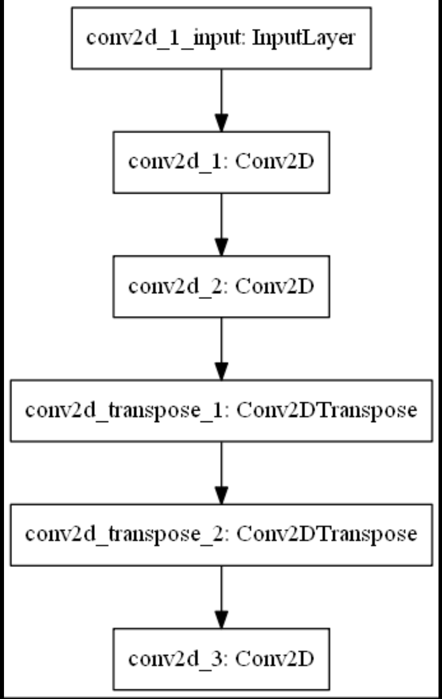


Denoising with AutoEncoder networks

Noise

- detection can be accomplished even on noisy signal, but sorting requires denoising
- we added artificial random noise and trained an AE network on it

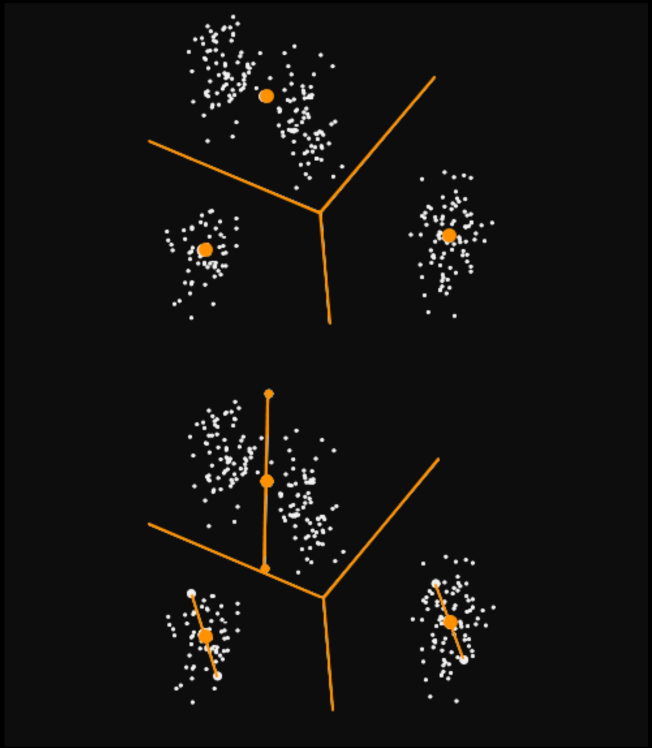




X-means and PCA-based sorting

About X-means

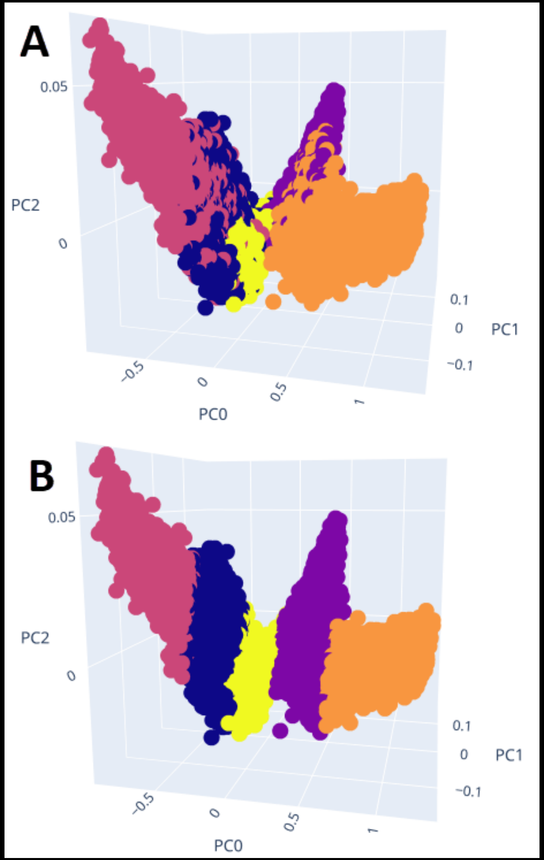
- similar to k-means, but tries to solve three main issues:
  - Poor scalability
  - The number of categories (k) has to be manually supplied in advance
  - Finishing prematurely because of local minima
- Runs like normal k-means, but occasionally splits centroids and evaluates the results



X-means: Extending K-means with Efficient Estimation of the Number of Clusters
Dan Pelleg, Andrew Moore

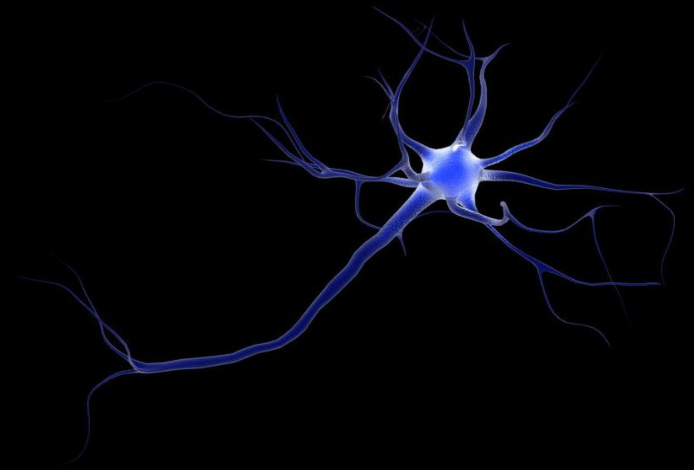
Classification

- simple, classical method: determining key features and measures of the waveform
- PCA is better by definition
- X-means always found the maximum set cluster number
- 71.33% accuracy in 10 dimensional space
    - A - correct
    - B - X-means guess



Conclusions

- 98.9% balanced classification accuracy
- RF is best with small training data
- AutoEncoder is suitable for denoising
- X-means can’t be used to find k



- future work:
  - real data
  - lstm, rnn