Unsupervised Automatic Spike Sorting

Daniel Sommerman, David Brody, Michael Gummelt
Department of Computer Science, Stanford University, Stanford, California

Introduction

Extracellular neurophysicological data recorded from a microelectrode is a mixture of the responses from the many surrounding cells. The responses are then clustered into similar shaped groups which represent different cells. Our project aims to take this analysis step which is usually done manually and automate it using machine learning.

Problem and Data Description

We are given extracellular recordings from 60 different channels. Each channel represents a data stream of voltage potentials at a specific microelectrode. A sample recording looks like the following:

The aim of the project is to cluster each spike into clusters which each share a similar shape. Example cluster centers could look like the following:

Model Definitions

Model Parameters
N = of datapoints per channel
CH = of channels
S = of points to smooth raw data over
T = Threshold for spikes
L = points to take from left of spike location R = points to take from right of spike location

Definitions
We begin with the raw waveform:

$$R = R^i \in R^N; i = 1, ..., CH$$

Smoothing over S points we define a new waveform where R^i_j is the jth datapoint in the ith channel:

$$W^i_j = \frac{1}{s}\sum_{s=-S/2}^{S/2} R^i_{j+s}$$

We then define W' as the zero mean normalized to the standard deviation on each channel of W. The next step is to separate out the possible spikes to cluster. We define the spike locations of interest as:

$$P = \{i; \frac{\partial}{\partial W} = 0 \cap (W'_1(P_i) > T \cup ... \cup W'_{CH}(P_i) > T) \cap i \in 1, ..., N\}$$

Peaks are then cut to ensure that they are a minimum of $\alpha = 15$ from each other. To form a single feature vector $X^{(i)}$ we take the surrounding points on each channel and concatentate them.

$$X^{(i)}_j = W'^{floor(i/(L+R+1))+1}_{mod(j,L+R+1)+P(i)}; i = 1, ..., |P|, j = 1, ..., (L + R + 1) * CH$$

X is the feature vector we use to then cluster on.

Pipeline

images/noexist.png

We paramaterize our feature vectors from the raw data with the following representations:

$$X = M(R; S, T, L, R)$$

We then use different methods to reduce the dimensionality of the feature vectors to improve performance and to emphasize more discrimanent features.

$$Red(X, N, M; X \in \mathbb{R}^N) = Y \in \mathbb{R}^M$$

Clustering then takes the dimension reduced feature vectors and assigns them to clusters which represent possible cells.

Dimensionality Reduction

Principle Component Analysis

images/noexist.png

This is a column!
Polynomial Fitting

Clustering

G-Means
Davies

Conclusion

Conclusions
- Achieved result close to graph-structure based link predictors
- Experiments show that success of prediction comes from having accurate univariate link potential
- Model preserves information
- Adding classifications of people does not significantly add to predictive power
  - May be because Email is poor indicator of useful class partitioning

Learnings
- Make sure features chosen have suffient mutual informtion with variables of interest

Future Work
- Try with features of email that may be more indicitive of future communications
- Try in domain that has more direct features of who a person may communicate with
  - Facebook information would allow for application to social network structure as well as indicitive features of each person including location, school, interests, etc.
- Test model in scenario where limited amounts of test data are observable