Shallow Transits

Deep Learning

Shay Zucker
Raja Giryes

Tel Aviv University

ISSI international team
Juan Cabrera
Biological Neuron
McCulloch-Pitts Neuron

The McCulloch-Pitts Neuron model is a simple computational model of a neuron. It consists of inputs, synaptic weights, a summing function, a non-linear activation function, and an output. The model can be represented by the following equation:

\[ y_k = \sigma \left( \sum w_k x_i + b_k \right) \]

Where:
- \( y_k \) is the output of neuron \( k \)
- \( w_k \) are the synaptic weights
- \( x_i \) are the inputs
- \( b_k \) is the bias
- \( \sigma \) is the non-linear activation function
Deep Neural Networks

Inspiration: human visual cortex
Training using Backpropagation

- Supervised learning: given examples with ground truth ('training set')
- Loss function (error quantification)
- Loss depends analytically on the synaptic weights
- Back-propagation of derivatives (chain rule) through layers
- Slowly update the synaptic weights (e.g. gradient descent, Metropolis-Hastings, etc.) to minimize loss
Deep Learning

• Comeback of neural networks (~2005)
• Depth made all the difference
• Depth (number of layers) allows significant abstraction
• Progress made possible by hardware
• Breaks records of performance
Deep Learning Revolution

- Computer vision
  - Scene parsing, face recognition, handwriting recognition, etc.
- Speech recognition
- Automatic machine translation
- Genomics (e.g. roles of non-coding DNA sequences)
- Drug discovery (e.g. predict metabolic fate of a molecule)
- Autonomous cars
- ... and many more
Essential Ingredients

- Large and comprehensive training set
- Deep network (many layers)
- Adequate network architecture
  - Convolutional neural networks (ConvNets)
  - Fully Convolutional networks (FCN)
  - Recurrent neural networks (RNN)
  - Residual networks (ResNets)
- Back-propagation scheme
Typical Task #1: Image Classification

Microsoft competition: dogs vs. cats (Kaggle dataset)
Design a code that will distinguish between cats and dogs

98.9% right! (Pierre Sermanet, 2014)
Typical Task #2: Image Denoising

Remez et al. 2017
Typical Task #3: Image Segmentation

Shelhamer et al. 2017
Relevant Tasks for PLATO

- Transit detection (‘classification’, #1)
- Detrending (‘denoising’, #2)
- Individual transit identification (‘segmentation’, #3)
- Estimates of false positive/negative rates
Feasibility Study

- Cadence 5 min
- Time span ~21 days
- (6144 samples)
- Training set:
  - 83333 lightcurves with transits
  - 83333 lightcurves with no transits
Feasibility Study

- GP Simulated red noise:
  \[ k(t_i - t_j) = A_s^2 \exp\left[ -\left( \frac{t_i - t_j}{\lambda_s} \right)^2 \right] + A_q^2 \exp\left[ -\frac{\sin^2[\pi(t_i - t_j)/T_q]}{2} - \left( \frac{t_i - t_j}{\lambda_q} \right)^2 \right] + A_w^2 \delta(t_i - t_j) \]
- \( A_s \sim 20 - 500 \) ppm, \( A_q \sim 2 - 500 \) ppm
- \( A_w = 140 \exp[0.2(M-M_{\text{min}})] \)
- \( M \sim 10 - 16 \)
- \( \lambda_s \sim 1 \) min – 10 hours, \( T_q \sim 10 - 500 \) hours
- \( \lambda_q \sim 16.6 - 500 \) hours
Feasibility Study

- Trapezoidal simulated transits:
  - \( P \sim 16.6 \text{ hours} - 4.2 \text{ days} \)
  - Depth \( \sim 10^{-3} - 10^{-4} \)
  - Duration \( \sim 30 \text{ min} - 3.3 \text{ hours} \)
Feasibility Study

• One-layer FCN for detection
• (actually not ‘deep’ learning)
• Applied on the Fourier amplitudes
• Testing:
  – 16667 lightcurves with transits
  – 16667 lightcurves with no transits
• 7% false positives, 12% false negatives
Feasibility Study

SNR proxy: \( \frac{d}{\sigma} \sqrt{\frac{w}{P}} \)
Feasibility Study
Feasibility Study
Conclusions

• Deep learning neural networks may be the way forward.
• They may achieve unprecedented results
• A fundamentally different approach