CS771 Project

How Close are ANNs to the Brain?

Shashi Kant Gupta and Syed Naveed

December 4, 2018

Indian Institute of Technology Kanpur

Contents

1. Why this Project?
2. Present Brain View of Neural Networks
3. Our Work
4. Conclusion

Why this Project?

Why this Project?

- To Understand Some Existing ANN Models

Why this Project?

- To Understand Some Existing ANN Models
- And Their Biological Plausibility

Why this Project?

- To Understand Some Existing ANN Models
- And Their Biological Plausibility
- To Bridge the Gap between Deep Learning and Neuroscience

Why this Project?

- To Understand Some Existing ANN Models
- And Their Biological Plausibility
- To Bridge the Gap between Deep Learning and Neuroscience
- With a Hope to Gear Toward More Scalable Learning Algorithms

Present Brain View of Neural Networks

Communication in Biological Neurons

- Sensors/ Receptors receives inputs as current and produces voltages in connecting neurons
- If the produced voltage is above a certain threshold it produces spikes which travel along the axon

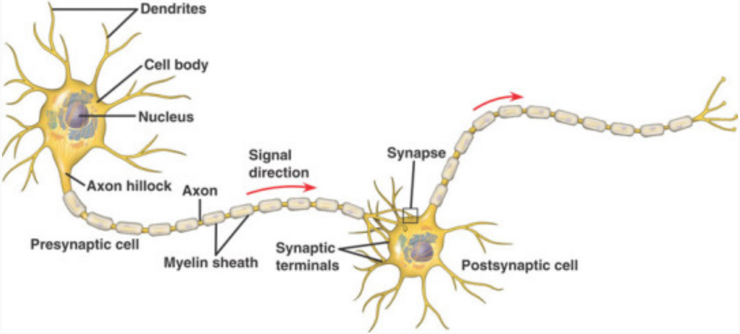

Figure 1: A drawing of biological neuron. Source: https://pmgbiology.com/

Communication in Biological Neurons

- Axons branches out and connects to dendrites of other neurons via synapses which produce voltages in them based on the spiking activity of initial neuron
- These synapses are **learnable** features (Weights) of neurons which changes according to some learning rule

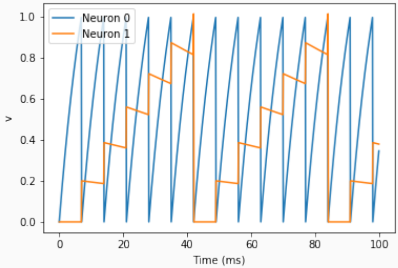

Figure 2: An example showing how spikes of one neuron produces spikes in another. Plot generated using Brian2 Simulator

- Synaptic weights changes if there are pre-synaptic spikes around a post-synaptic spike
- The change is positive if the post-synaptic spike is after the pre-synaptic spike and vice-versa

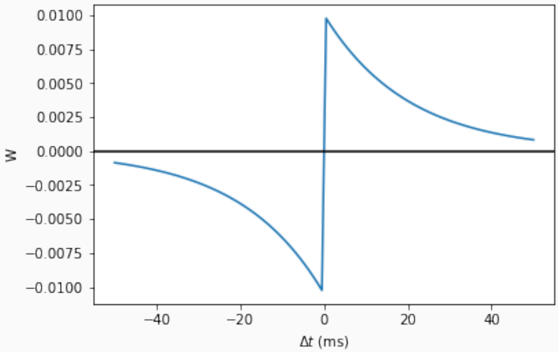

Figure 3: Visualisation of STDP

Analogy

- Assumes precise timing of spikes doesn't matter. And f(Wx) gives the firing rate of neurons
- Learning is through backpropagation of errors

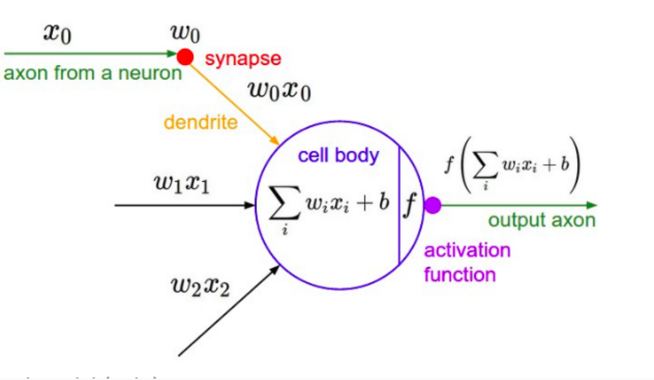

Figure 4: Artificial Neuron. Source: http://cs231n.github.io

Issues

- Backprop is not biologically plausible.
- Feedback path will need exact weights of feedforward path to perform backprop
- BP of gradients is also linear which is not plausible
- Biological neurons communicate via Spikes
- Loss of information of precise timing of Spikes.

Research Going On ...

- Many research going on Spiking Neural Networks to resemble Brain Networks but not much from the viewpoint of Machine Learning Algorithm.
- **Peter U. Diehl**, Implemented an SNN architecture based on STDP to get 95% accuracy on the MNIST using the Brian2 Simulator
- **Yoshua Bengio**, Shown a Machine Learning Interpretation of STDP rule and a learning rule governed by it using a latent variable setup for post-synaptic voltage (more..)

More on Yoshua Bengio Paper: Towards Biologically Plausible Deep Learning

- Assumed STDP rule: $\Delta W_{ij} \propto S_i \Delta V_{ij}$; S - pre-spike, V - post-synapse voltage
- Can be related to an **SGD** update on Objective J if: $\Delta V_{ij} \approx \partial J / V_{ij}$
- Taking post-synapse volatage as a latent variable h, he proposed an EM learning algorithm

$$log(p(x;\theta)) \geq E_{q(H|x)}[log(p(x,H;\theta))] \tag{1}$$

$$J = log(p(x,h;\theta)) + regularizer \tag{2}$$

Our Work

Implementation of CNN and RNN model on MNIST Dataset

This was carried on to familiarise ourselves to some of the present ANN models as well as to compare different ANN structures. Below is a brief overview of what we worked on:

- Modelling of a **CNN** architecture to get an accuracy upto **99.49%**

Implementation of CNN and RNN model on MNIST Dataset

This was carried on to familiarise ourselves to some of the present ANN models as well as to compare different ANN structures. Below is a brief overview of what we worked on:

- Modelling of a **CNN** architecture to get an accuracy upto **99.49%**
- **Ensembling** on 5 different CNN models helped us in achieving **99.60%** (Stocking One)

Implementation of CNN and RNN model on MNIST Dataset

This was carried on to familiarise ourselves to some of the present ANN models as well as to compare different ANN structures. Below is a brief overview of what we worked on:

- Modelling of a **CNN** architecture to get an accuracy upto **99.49%**
- **Ensembling** on 5 different CNN models helped us in achieving **99.60%** (Stocking One)
- Preprocessing the dataset to two different time series data to learn an **RNN** Model

Implementation of CNN and RNN model on MNIST Dataset

This was carried on to familiarise ourselves to some of the present ANN models as well as to compare different ANN structures. Below is a brief overview of what we worked on:

- Modelling of a **CNN** architecture to get an accuracy upto **99.49%**
- **Ensembling** on 5 different CNN models helped us in achieving **99.60%** (Stocking One)
- Preprocessing the dataset to two different time series data to learn an **RNN** Model
- Both gave us an accuracy around **98.9%** (Stocking One)

Implementation of CNN and RNN model on MNIST Dataset

This was carried on to familiarise ourselves to some of the present ANN models as well as to compare different ANN structures. Below is a brief overview of what we worked on:

- Modelling of a **CNN** architecture to get an accuracy upto **99.49%**
- **Ensembling** on 5 different CNN models helped us in achieving **99.60%** (Stocking One)
- Preprocessing the dataset to two different time series data to learn an **RNN** Model
- Both gave us an accuracy around **98.9%** (Stocking One)
- Their ensemling gave us a dramatic improvement upto **99.38%** (Stocking One)

Explaining How Conventional NN Carries Info about Firing Rates

Here we tried explaining how the present model can be realized as spikes and they really do carry the information of firing rate forward in the network!

- Taking the feature x_i of observed variable $\mathbf{x}$ as the produced current at time t_i. The produced voltage can be assumed to be $V_i = V_{i-1} + W_i x_i$ at time $t = t_i$.

Explaining How Conventional NN Carries Info about Firing Rates

Here we tried explaining how the present model can be realized as spikes and they really do carry the information of firing rate forward in the network!

- Taking the feature x_i of observed variable $\mathbf{x}$ as the produced current at time t_i. The produced voltage can be assumed to be $V_i = V_{i-1} + W_i x_i$ at time $t = t_i$.
- Threshold Voltage at which spikes occurs, to be V_{th}

Explaining How Conventional NN Carries Info about Firing Rates

Here we tried explaining how the present model can be realized as spikes and they really do carry the information of firing rate forward in the network!

- Taking the feature x_i of observed variable $\mathbf{x}$ as the produced current at time t_i. The produced voltage can be assumed to be $V_i = V_{i-1} + W_i x_i$ at time $t = t_i$.
- Threshold Voltage at which spikes occurs, to be V_{th}
- So, no of spikes $\approx \sum W_i x_i / V_{th} \propto$ firing rate for some Δt over the full observations.

Explaining How Conventional NN Carries Info about Firing Rates

Here we tried explaining how the present model can be realized as spikes and they really do carry the information of firing rate forward in the network!

- Taking the feature x_i of observed variable $\mathbf{x}$ as the produced current at time t_i. The produced voltage can be assumed to be $V_i = V_{i-1} + W_i x_i$ at time $t = t_i$.
- Threshold Voltage at which spikes occurs, to be V_{th}
- So, no of spikes $\approx \sum W_i x_i / V_{th} \propto$ firing rate for some Δt over the full observations.
- Infact we can also show that the generated spikes carry the same info as that of a RELU activation to next neuron. Since the increase in the voltage of next neuron is only when spikes occur in the pre-synapse neuron.

Explaining How Conventional NN Carries Info about Firing Rates

Here we tried explaining how the present model can be realized as spikes and they really do carry the information of firing rate forward in the network!

- Taking the feature x_i of observed variable $\mathbf{x}$ as the produced current at time t_i. The produced voltage can be assumed to be $V_i = V_{i-1} + W_i x_i$ at time $t = t_i$.
- Threshold Voltage at which spikes occurs, to be V_{th}
- So, no of spikes $\approx \sum W_i x_i / V_{th} \propto$ firing rate for some Δt over the full observations.
- Infact we can also show that the generated spikes carry the same info as that of a RELU activation to next neuron. Since the increase in the voltage of next neuron is only when spikes occur in the pre-synapse neuron.
- Increase in Voltage of post-synapse neuron = (no of pre-spikes)*W2 $\approx W2_{ij} * (\sum W1_{jk} x_k / V_{th})$

Can We Carry Precise Info of Spikes Time?

Removed because of ongoing work and privacy reason

Conclusion

Conclusion

- Although the ANN has really progressed a lot but its still very far from Neuroscience perspective.
- Work has been going to bridge the gap between the two and to achieve a scalable learning algorithm.
- We tried to add a little contribution to it.
- Unfortunately right now we don’t have any empirical validity for this we are still working on its implementation part.

Thank You!

- We would like give our thanks to Prof Piyush Rai for giving us this opportunity to work on this project
- Lastly, thanks to all of you who are present here for listening to our presentation