

Reading List

Daniel Saunders

September 24, 2018

Note: Work in progress. Some dates before mid-July 2018 are approximate.

1 Books

1.1 Read

- *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*
Read date: 2018-08-09
<https://mitpress.mit.edu/books/theoretical-neuroscience>
Synopsis: The mathematical study of neural systems is attacked from three separate perspectives: (1) neural encoding and decoding; i.e., the study of how natural stimuli are converted into neural responses or action potentials and vice versa, (2) neurons and neural circuits, from single-compartmental and multi-compartmental spiking models of neurons to rate-based networks of neural units, and (3) adaption and learning, including biologically realistic notions of synaptic plasticity and forms of conditioning to abstract, computationally motivated learning and inference. The relationship of real neural data with mathematical models is emphasized. Ideas from mathematical modeling, probability and statistics, information theory, and machine learning are all brought to bear on problems in the neurosciences.
- *Computer Age Statistical Inference*
Read date: 2018-07-01
<https://web.stanford.edu/~hastie/CASI/>
Synopsis: The history of the development of statistical methods and justification of those methods is divided into three periods: classic statistical inference, early computer-age methods, and 21st century topics. Each statistical method is discussed in terms of their frequentist, Bayesian, or Fisherian justification(s), and examples with real datasets are common throughout. The authors describe how the development of statistical algorithms is out-pacing their inferential justification thanks to powerful and cheap computation and the collection of massive, interesting datasets.
- *Deep Learning*
Read date: 2018-06-01
<http://www.deeplearningbook.org/>
Synopsis: The nascent field of deep learning is described starting with machine learning basics, working through established deep learning concepts (multilayer networks, convolutional networks, recurrent networks, training strategies, and applications), and culminating in an in-depth discussion of research topics. Basic calculus, linear algebra, probability, and statistics are all that's needed to understand most of the material.
- *The Computer and the Brain*
Read date: 2018-05-01
<https://dl.acm.org/citation.cfm?id=578873>
Synopsis: The building blocks and working mechanisms of computers and biological brains are compared and contrasted. Von Neumann attempts to explain some of the workings of the nervous system from the point of view of a mathematician.

- *Your Brain is a Time Machine*

Read date: 2018-04-01

<http://books.wwnorton.com/books/Your-Brain-Is-a-Time-Machine/>

Synopsis: It is argued that the human brain is a complex system that not only tells time, but creates it, constructing our sense of chronological movement and enabling mental time travel; i.e., simulations of past and future events. These functions were and are essential for the evolution of the human race, allowing us to anticipate the need for tools and farming.

1.2 Reading

- *Quantum Computation and Quantum Information: 10th Anniversary Edition*
<https://dl.acm.org/citation.cfm?id+AD0-1972505>
- *The Elements of Statistical Learning*
<https://www.springer.com/us/book/9780387848570>
- *The Book of Why: The New Science of Cause and Effect*
<https://www.basicbooks.com/titles/judea-pearl/the-book-of-why/9780465097609/>
- *Causality: Models, Reasoning, and Inference*
<http://bayes.cs.ucla.edu/BOOK-2K/>

2 Papers

2.1 Read

- *Gradient Descent for Spiking Neural Networks*
Read date: 2018-09-24
<https://arxiv.org/pdf/1706.04698.pdf>
Synopsis: Most studies of neural computation are based on network models of rate-based neurons that process analog input, despite the fact that information processing in the brain is predominantly done by dynamic neurons that produce spikes. Research in spike-based computation is impeded by the lack of an efficient supervised learning algorithm for spiking neural networks (SNNs). The authors present a gradient descent method for optimizing SNN models by introducing a differentiable formulation of SNNs and deriving the exact gradient calculation. They train recurrent SNNs for two dynamic tasks: one requiring millisecond time scale spike-based interactions for efficient information encoding, and a delayed-memory XOR task over extended time (on the order of seconds). The method is able to optimize SNN dynamics on the time scale of individual spikes as well as behavior. The result is a general purpose supervised learning algorithm for SNNs.
- *Supervised learning in spiking neural networks with FORCE training*
Read date: 2018-09-23
<https://www.nature.com/articles/s41467-017-01827-3.pdf>
Synopsis: FORCE training, previously introduced to trained networks of rate-based neurons, is used to train spiking neural networks (SNNs) to mimic dynamical systems, classify inputs, store discrete sequences that correspond to notes in a song, and mimic two biologically motivated model circuits. The first is used to reproduce songbird singing in the adult zebra finch, and the second is trained to store and replay a movie scene, which requires a high-dimensional temporal signal (HDTs), jointly trained or fixed, in order to successfully learn. FORCE-trained networks reproduce behaviors comparable in complexity to those of their biologically inspired circuits and yield information not easily obtainable with methods such as pharmacological manipulations or spike time statistics.
- *Causal inference in statistics: an overview*
Read date: 2018-09-13
<http://ftp.cs.ucla.edu/pub/stat+AF8-ser/r350.pdf>
Synopsis: This review paper presents recent advances in causal inference and emphasizes paradigmatic

shifts needed to move from traditional statistical analysis to causal analysis. Emphasis is also placed on assumptions underlying all of causal inference, the language used in formulating these assumptions, the conditional nature of causal and counterfactual claims, and methods developed for the assessment of these claims. The general theory of causation based on the Structural Causal Model (SCM) provides a mathematical foundation for the analysis of causes and counterfactuals. The paper surveys the development of tools for inferring (from data and assumptions) answers to three types of causal queries: (1) queries about the effects of intervention (causal effects / policy evaluation), (2) queries about probabilities of counterfactuals (assessment of regret, attribution, or causes of effects), and (3) queries about direct and indirect effects (mediation). The paper also describes the relationships between the SCM and potential-outcome frameworks.

- *Task-Driven Convolutional Recurrent Models of the Visual System*
Read date: 2018-07-19
<https://arxiv.org/abs/1807.00053>
Synopsis: While convolutional neural networks are SOTA for object classification and accurately model time-averaged visual system neural responses, they do not have two important features of biological visual systems: (1) local recurrence within cortical areas, and (2) long-range feedback from downstream to upstream areas. Standard vanilla RNN and LSTM do not perform well in deep CNNs on ImageNET, but custom cells using bypassing and gating do. Automated architecture search is used to identify new local recurrent cells and long-range connections for the ImageNET task, which well-explain the neural dynamics in the primate visual system.
- *Putting a bug in ML: The moth olfactory network learns to read MNIST*
Read date: 2018-07-12
<https://arxiv.org/abs/1802.05405>
Synopsis: The moth olfactory network model of Biological Mechanisms for Learning: A Computational Model... is used to classify the MNIST digits using a very small number of samples per digit class (1-20).
- *Relational inductive biases, deep learning, and graph networks*
Read date: 2018-06-15
<https://arxiv.org/abs/1806.01261>
Synopsis: Many defining characteristics of human intelligence remain out of the reach of current popular artificial intelligence techniques in particular, generalizing beyond one's experiences. It is argued that combinatorial generalization should be a priority for AI research to achieve human-level abilities, and structured representations and computations over these are key to realizing this. The authors reject the choice between hand-engineering and end-to-end learning, instead advocating an approach benefiting from their complementary strengths. They show how relational inductive biases can be used in deep learning to facilitate learning about entities, relations, and compositions thereof. The graph network is presented, generalizing and extending previous approaches to machine learning on graphs.
- *SuperSpike: Supervised Learning in Multilayer Spiking Neural Networks*
Read date: 2018-06-08
<https://www.mitpressjournals.org/doi/abs/10.1162/neco+AF8-a+AF8-01086>
Synopsis: This paper provides a possible answer to the question: How do neural circuits learn and compute in vivo, and how can such abilities be implemented in artificial spiking circuits in silico? The problem of supervised learning in temporally coded multilayer spiking networks is studied. A surrogate gradient approach is used to derive the SuperSpike learning rule, capable of training multilayer networks of IF neurons to do nonlinear computations on spatiotemporal inputs. The performance of the learning rule under different credit assignment strategies for propagating error signals to hidden units is compared (uniform, symmetric, and random feedback).
- *Differentiable plasticity: training plastic neural networks with backpropagation*
Read date: 2018-05-17
<https://arxiv.org/abs/1804.02464>
Synopsis: A simple solution to the problem of meta-learning is proposed, taking inspiration from

learning in biological brains: synaptic plasticity (just like connection weights) can be optimized via gradient descent in large recurrent networks with Hebbian plasticity. The meta-learning technique is demonstrated on three simple tasks: memorizing and reconstructing high-dimensional natural images, the Omniglot task (a generic one-shot learning task), and a maze exploration reinforcement learning problem.

- *Training Deep Spiking Neural Networks Using Backpropagation*
Read date: 2018-05-01
<https://www.frontiersin.org/articles/10.3389/fnins.2016.00508/full>
Synopsis: Spiking neural networks may improve the latency and energy efficiency of deep neural networks using event-based computation, though their training is difficult due to their non-differentiability. The authors treat neuron membrane potentials as differentiable signals, where spike time discontinuities are considered noise, effectively enabling backpropagation on spike signals and membrane potentials. The technique is demonstrated on the MNIST and N-MNIST (neuromorphic) datasets, where it is shown that equivalent accuracy can be achieved with much less computation.
- *Biological Mechanisms for Learning: A Computational Model of Olfactory Learning in the Manduca sexta Moth, with Applications to Neural Nets*
Read date: 2018-04-15
<https://arxiv.org/abs/1802.02678>
Synopsis: A model of the insect olfactory system (in particular, the moth) is built using integrate-and-fire neurons is tuned to reproduce experimental in vivo firing rate data. The model is trained to learn new odors using very few data samples.
- *Reinforcement Learning Through Modulation of Spike-Timing-Dependent Synaptic Plasticity*
Read date: 2018-04-15
<https://ieeexplore.ieee.org/document/6796089/>
Synopsis: Modulation of STDP by a global reward signal leads to reinforcement learning in spiking neural networks. Learning rules are analytically derived by applying a reinforcement learning algorithm to a stochastic response model of spiking neurons. Two simplified reward-modulated learning rules are shown to be effective in simulations of IF neuron networks. The first rule is a direct extension of standard STDP (modulated STDP), and the second involves an eligibility trace for each synapse that tracks a decaying memory of pre- and post-synaptic spiking activity (modulated STDP with eligibility trace).
- *A Model for Real-Time Computation in Generic Neural Microcircuits*
Read date: 2018-04-01
<https://papers.nips.cc/paper/2307-a-model-for-real-time-computation-in-generic-neural-microcircuits.pdf>
Synopsis: Liquid state machines (LSMs) are introduced and are motivated from the point of view of the anytime computing / real-time computing paradigms inspired by neural computation. In particular, a simulation of a small network of heterogeneous LIF neurons are used to filter input signals $u(t)$ into a liquid state $x(t)$, and a small set of linear read-out filters are optimized to output a target time series $y(t)$. A non-Turing (that is, continuous in time and real-valued) theory of computation is developed with the LSM, with the result that a sufficiently large / complex found or evolved generic circuit will tend to have sufficient computational power for any given real-valued, parallel real-time computing task. An LSM with a small generic neural circuit as the computation reservoir is shown to achieve SOTA results on a dataset of 500 (300 train / 200 test) audio examples of the spoken digits 0-9, with several desirable properties (any-time outputs).
- *Spiking allows neurons to estimate their causal effect*
Read date: 2018-04-01
<https://www.biorxiv.org/content/early/2018/01/25/253351>
Synopsis: Regression discontinuity design (a popular causal technique from economics) is used in a new synaptic learning rule such that neurons may estimate their causal effect on task performance.
- *Unsupervised Feature Learning With Winner-Takes-All Based STDP*
Read date: 2018-04-01

<https://www.frontiersin.org/articles/10.3389/fncom.2018.00024/full>

Synopsis: Spike-timing-dependent plasticity (STDP) is used to learn image features from the MNIST, ETH80, CIFAR-10, and STL-10 datasets, which are subsequently used for classification. The authors show an equivalence between rank order coding LIF neurons and ReLUs when applied to non-temporal data. A binary STDP rule is derived and used to perform batched training on image data. A winner-takes-all (WTA) mechanism selects the most relevant patches to learn from among the spatial dimensions, and a feature-wise normalization is used to maintain homeostatic activity. Ultimately, their networks are able to learn multi-layer convolutional sparse features.

- *Bayesian GAN*

Read date: 2018-03-15

<https://arxiv.org/abs/1705.09558>

Synopsis: A practical Bayesian formulation for unsupervised and semi-supervised learning with GANs is developed. Stochastic gradient Hamiltonian Monte Carlo is used to marginalize the weights of the generator and discriminator networks. This approach removes the need for the typical GAN training interventions such as feature matching or minibatch discrimination. It is also able to avoid mode-collapse, produces interpretable and diverse generated samples, and achieves SOTA quantitative results for semi-supervised learning on the SVHN, CelebA, and CIFAR-10 datasets.

- *Biologically inspired load balancing mechanism in neocortical competitive learning*

Read date: 2018-03-01

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3949291/>

Synopsis: The authors present a simulation of a population of 1000 LIF neurons with layer 5 Martinotti cells (MC) with delayed self-inhibition and layer 5 large basket cell with local Mexican hat-shaped inhibition, as well as STDP learning of the synapses which are connected with distance-dependent randomness between neurons. Input to the network are random bit vectors. The results show that the network is able to organize into a few large (or many small and overlapping) clusters, which compete between themselves yet synchronize within themselves. Various kinds of cluster analysis are applied to the connectivity and activity of the trained network.

- *Unsupervised learning of digit recognition using spike-timing-dependent plasticity*

Read date: 2017-03-01

<https://www.frontiersin.org/articles/10.3389/fncom.2015.00099/full>

Synopsis: A spiking neural network model is described and used to classify the MNIST handwritten digits. Spike-timing-dependent plasticity is used to update synapse weights, inhibition is used to create competition between neurons filtering the input, and an adaptive, homeostatic mechanism is used to adjust sensitivity to different input intensities.