Fall 2023:

Basics of supervised learning and geometry of ReLU networks:

- M. Nielsen, *Neural networks and deep learning*
- Grigsby, Lindsey: *On transversality of bent hyperplane arrangements and the topological expressiveness of ReLU neural networks*
- M. Masden: *Algorithmic determination of the combinatorial structure of the linear regions of ReLU networks*
- Grigsby, Lindsey, Meyerhoff, Wu: *Functional dimension of ReLU networks*
- Grigsby, Lindsey, Rolnick: *Hidden symmetries of ReLU networks*
- Baraniuk: *Affine spline insights into deep learning*
- Baraniuk: *Geometry of deep networks: power diagram subdivision*

Generalization & Statistical Notions of Complexity: VC dimension, PAC learning, Rademacher complexity, covering numbers, Generalization

- Sontag, E.: *VC dimension of neural networks*
- Kearns, Vazirani: *An introduction to computational learning theory*
- *Learnability and VC dimension*
- Shai-Shwartz, Ben-David: Part I of *Understanding Machine Learning: Theory and Algorithms*
- Neyshabur et.al.: *Exploring generalization in deep learning*
- Bartlett, P. et. al.: "*Nearly tight VC dimension bounds*"
- Chen, Klivans, Meka: *Learning Deep ReLU networks is fixed-parameter tractable*
- Ovchinnikov, Sergei: *Max-min representations of piecewise-linear functions*

Neural tangent kernel and Gaussian processes:

- See references and code at this github page

Tropical geometry and ReLU networks:

- Zhang, L. et.al.: *Tropical geometry of deep neural networks*
- Haase, C. et.al.: *Lower bounds on the depth of integral ReLU neural networks via lattice polytopes*

Implicit regularization and geometry of loss landscape:

- Y. Cooper: *Loss landscape of overparameterized networks*
- Q. Nguyen: *On connected sublevel sets in deep learning*
- *Visualizing mode connectivity* blog post
- *How to escape saddle points efficiently* blog post
- Damian, Lee, Ma: *Label noise SGD provably prefers flat global minimizers*
- Bruna, Trager, et.al.: "*Gradient dynamics of shallow univariate ReLU Networks*"
- Vardi, G. et.al.: *On the effective number of linear regions in shallow univariate ReLU networks: convergence guarantees and implicit bias*

Attention & Transformer networks:

- Olah, C.: *A mathematical framework for transformer circuits*
- Vaswani, A. et.al.: *Attention is all you need*

If you are new to the subject, here is an augmented blurb I sent to UT, Austin prior to a couple of talks I gave there in April:

Pre-talk resources:
The main things it would be nice for people to know on the machine learning side are:

- Basic idea behind supervised learning problems, especially regression tasks
- What a feedforward neural network is

A good resource for this is: Chapter 8 and the beginning of Chapter 1 of the attached MIT Lecture Notes (from an open courseware course). I also really like the exposition in the first chapter of Michael Nielsen's on-line book Neural networks and deep learning (with the caveat that Nielsen focuses on *sigmoid* neurons, and I will focus on *ReLU* neurons, which is what modern practitioners use for regression tasks)

Another option for background that could be useful (completely up to the taste of whoever is giving the pre-talk) is the basics of affine hyperplane arrangements, polyhedral sets, and transversality, since this enters the mathematical story of neural networks early on: the presenter could follow the first couple of background sections of my first paper with Kathryn Lindsey. The most important notion is what it means for an affine hyperplane arrangement to be *generic* and what this has to do with transversality (submanifolds intersect with the expected codimension). (We extend this notion to *bent hyperplane arrangements*, which is what one naturally encounters when studying ReLU neural networks).

You might also be interested in this paper of Marissa Masden, which greatly improves our definition of "transversal," and allows her to build a cubical complex (and accompanying code!) that extracts the homological invariants of decision boundaries and decision regions for binary classification tasks.