Statistics 37793 — Topics in Deep Learning: Discriminative Models

Project Paper Lists

- **Batch normalization**
  - What does BatchNorm do? [IS15, STIM19, BGSW18, MBRB18, FSM20]
  - Theory in special cases: [KDL+18, YPR+19, LWSP19]
  - Removing BatchNorm: [ZDM19, DS20]
  - **Possible projects:** Study the effect of BatchNorm on quantities related to generalization (trace of Hessian, etc.); study how hyperparameter choices change with/without BatchNorm.
- **Evolution of the NTK**
  - Neural tangent hierarchy: [HY19]
  - When is the NTK constant?: [JGH20, WGL+20, LZB20]
  - Empirical study of the NTK change during training: [LBD+20, FDP+20]
  - Package for computing the NTK: `https://github.com/google/neural-tangents`
  - **Possible projects:** Study how NTK dynamics change with network width; study effect of optimization hyperparameters (momentum, weight decay, label smoothing) on NTK change.
- **Implicit bias of SGD**
  - Flat and sharp minima: [KMN+17, DPBB17]
  - Batch size, learning rate, and gradient variance: [SKYL18, SL18, GDG+18, JKA+18]
  - Theory in special cases: [SHN+18, GLSS19, CB20, LL20, WLLM20, WS19]
  - **Possible projects:** Optimize hyperparameters for SGD on a new dataset; study how quantities related to generalization (gradient variance, trace of Hessian, etc.) vary over training.
- **Data augmentation**
  - Standard augmentation techniques: [Bis95, SHK+14, DT17, ZCDLP18]
  - Learning augmentation schedules: [CZM+19, CZSL19, HLS+19]
  - Augmentations for robustness: [HMC+20]
  - Theoretical analyses: [RFC+19, WZVR20, HS20]
  - **Possible projects:** Study the effect of augmentation on average-case robustness; compare representations learned with and without data augmentation.
- **Average-case robustness (distribution shift)**
  - CIFAR-10-C / ImageNet-C datasets: [HD19] (code at `https://github.com/hendrycks/robustness`)
  - Connecting average- and worst-case robustness: [FGCC19]
  - Shape vs. texture bias: [GRM+19] (code at `https://github.com/rgeirhos/texture-vs-shape`)
  - Many different robustness tasks: [HBM+20]
  - Distribution shift between different test sets for ImageNet and CIFAR-10: [RRSS19]
  - **Possible projects:** Try to improve performance on CIFAR-10-C; test a new defense technique.
- **Worst-case robustness (adversarial examples)**
  - Original paper introducing adversarial examples: [GSS15]
  - PGD attack and adversarial training: [MMS+19]
  - Breaking published defenses: [ACW18, TCBM20]
  - Best practices for evaluation: [CAP+19]
  - Certified defense via randomized smoothing: [CRK19]
  - Attacks outside L_p: [KSH+20]
  - **Possible projects:** Evaluate a published defense; test a new defense technique.

References

[ACW18] Anish Athalye, Nicholas Carlini, and David Wagner. Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples, 2018.

[BGSW18] Johan Bjorck, Carla Gomes, Bart Selman, and Kilian Q. Weinberger. Understanding batch normalization, 2018.

[Bis95] Chris M Bishop. Training with noise is equivalent to tikhonov regularization. *Neural computation*, 7(1):108–116, 1995.

[CAP+19] Nicholas Carlini, Anish Athalye, Nicolas Papernot, Wieland Brendel, Jonas Rauber, Dimitris Tsipras, Ian Goodfellow, Aleksander Madry, and Alexey Kurakin. On evaluating adversarial robustness, 2019.

[CB20] Lenaic Chizat and Francis Bach. Implicit bias of gradient descent for wide two-layer neural networks trained with the logistic loss, 2020.

[CRK19] Jeremy M Cohen, Elan Rosenfeld, and J. Zico Kolter. Certified adversarial robustness via randomized smoothing, 2019.

[CZM+19] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc V. Le. Autoaugment: Learning augmentation policies from data, 2019.

[CZSL19] Ekin D. Cubuk, Barret Zoph, Jonathon Shlens, and Quoc V. Le. Randaugment: Practical automated data augmentation with a reduced search space, 2019.

[DPBB17] Laurent Dinh, Razvan Pascanu, Samy Bengio, and Yoshua Bengio. Sharp minima can generalize for deep nets, 2017.

[DS20] Soham De and Samuel L. Smith. Batch normalization biases residual blocks towards the identity function in deep networks, 2020.

[DT17] Terrance DeVries and Graham W. Taylor. Improved regularization of convolutional neural networks with cutout, 2017.

[FDP+20] Stanislav Fort, Gintare Karolina Dziugaite, Mansheej Paul, Sepideh Kharaghani, Daniel M. Roy, and Surya Ganguli. Deep learning versus kernel learning: an empirical study of loss landscape geometry and the time evolution of the neural tangent kernel, 2020.

[FGCC19] Nic Ford, Justin Gilmer, Nicolas Carlini, and Dogus Cubuk. Adversarial examples are a natural consequence of test error in noise, 2019.

[FSM20] Jonathan Frankle, David J. Schwab, and Ari S. Morcos. Training batchnorm and only batchnorm: On the expressive power of random features in cnns, 2020.

[GDG+18] Priya Goyal, Piotr Dollr, Ross Girshick, Pieter Noordhuis, Lukasz Wesolowski, Aapo Kyrola, Andrew Tulloch, Yangqing Jia, and Kaiming He. Accurate, large minibatch sgd: Training imagenet in 1 hour, 2018.

[GLSS19] Suriya Gunasekar, Jason Lee, Daniel Soudry, and Nathan Srebro. Implicit bias of gradient descent on linear convolutional networks, 2019.

[GRM+19] Robert Geirhos, Patricia Rubisch, Claudio Michaelis, Matthias Bethge, Felix A. Wichmann, and Wieland Brendel. Imagenet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. In *International Conference on Learning Representations*, 2019.

[GSS15] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples, 2015.

[HBM+20] Dan Hendrycks, Steven Basart, Norman Mu, Saurav Kadavath, Frank Wang, Evan Dorundo, Rahul Desai, Tyler Zhu, Samyak Parajuli, Mike Guo, Dawn Song, Jacob Steinhardt, and Justin Gilmer. The many faces of robustness: A critical analysis of out-of-distribution generalization, 2020.

[HD19] Dan Hendrycks and Thomas Dietterich. Benchmarking neural network robustness to common corruptions and perturbations, 2019.

[HLS+19] Daniel Ho, Eric Liang, Ion Stoica, Pieter Abbeel, and Xi Chen. Population based augmentation: Efficient learning of augmentation policy schedules, 2019.

[HMC+20] Dan Hendrycks, Norman Mu, Ekin D. Cubuk, Barret Zoph, Justin Gilmer, and Balaji Lakshminarayanan. Augmix: A simple data processing method to improve robustness and uncertainty, 2020.

[HS20] Boris Hanin and Yi Sun. Data augmentation as stochastic optimization, 2020.

[HY19] Jiaoyang Huang and Horng-Tzer Yau. Dynamics of deep neural networks and neural tangent hierarchy, 2019.

[IS15] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift, 2015.

[JGH20] Arthur Jacot, Franck Gabriel, and Clment Hongler. Neural tangent kernel: Convergence and generalization in neural networks, 2020.

[JKA+18] Stanisaw Jastrzbski, Zachary Kenton, Devansh Arpit, Nicolas Ballas, Asja Fischer, Yoshua Bengio, and Amos Storkey. Three factors influencing minima in sgd, 2018.

[KDL+18] Jonas Kohler, Hadi Daneshmand, Aurelien Lucchi, Ming Zhou, Klaus Neymeyr, and Thomas Hofmann. Exponential convergence rates for batch normalization: The power of length-direction decoupling in non-convex optimization, 2018.

[KMN+17] Nitish Shirish Keskar, Dheevatsa Mudigere, Jorge Nocedal, Mikhail Smelyanskiy, and Ping Tak Peter Tang. On large-batch training for deep learning: Generalization gap and sharp minima, 2017.

[KSH+20] Daniel Kang, Yi Sun, Dan Hendrycks, Tom Brown, and Jacob Steinhardt. Testing robustness against unforeseen adversaries, 2020.

[LBD+20] Aitor Lewkowycz, Yasaman Bahri, Ethan Dyer, Jascha Sohl-Dickstein, and Guy Gur-Ari. The large learning rate phase of deep learning: the catapult mechanism, 2020.

[LL20] Kaifeng Lyu and Jian Li. Gradient descent maximizes the margin of homogeneous neural networks, 2020.

[LWSP19] Ping Luo, Xinjiang Wang, Wenqi Shao, and Zhanglin Peng. Towards understanding regularization in batch normalization, 2019.

[LZB20] Chaoyue Liu, Libin Zhu, and Mikhail Belkin. On the linearity of large non-linear models: when and why the tangent kernel is constant, 2020.

[MBRB18] Ari S. Morcos, David G. T. Barrett, Neil C. Rabinowitz, and Matthew Botvinick. On the importance of single directions for generalization, 2018.

[MMS+19] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks, 2019.

[RFC+19] Shashank Rajput, Zhili Feng, Zachary Charles, Po-Ling Loh, and Dimitris Papailiopoulos. Does data augmentation lead to positive margin?, 2019.

[RRSS19] Benjamin Recht, Rebecca Roelofs, Ludwig Schmidt, and Vaishaal Shankar. Do imagenet classifiers generalize to imagenet?, 2019.

[SHK+14] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.

[SHN+18] Daniel Soudry, Elad Hoffer, Mor Shpigel Nacson, Suriya Gunasekar, and Nathan Srebro. The implicit bias of gradient descent on separable data, 2018.

[SKYL18] Samuel L. Smith, Pieter-Jan Kindermans, Chris Ying, and Quoc V. Le. Don't decay the learning rate, increase the batch size, 2018.

[SL18] Samuel L. Smith and Quoc V. Le. A bayesian perspective on generalization and stochastic gradient descent, 2018.

[STIM19] Shibani Santurkar, Dimitris Tsipras, Andrew Ilyas, and Aleksander Madry. How does batch normalization help optimization?, 2019.

[TCBM20] Florian Tramer, Nicholas Carlini, Wieland Brendel, and Aleksander Madry. On adaptive attacks to adversarial example defenses, 2020.

[WGL+20] Blake Woodworth, Suriya Gunasekar, Jason D. Lee, Edward Moroshko, Pedro Savarese, Itay Golan, Daniel Soudry, and Nathan Srebro. Kernel and rich regimes in overparametrized models, 2020.

[WLLM20] Colin Wei, Jason D. Lee, Qiang Liu, and Tengyu Ma. Regularization matters: Generalization and optimization of neural nets v.s. their induced kernel, 2020.

[WS19] Mingwei Wei and David J Schwab. How noise affects the hessian spectrum in overparameterized neural networks, 2019.

[WZVR20] Sen Wu, Hongyang R. Zhang, Gregory Valiant, and Christopher R. On the generalization effects of linear transformations in data augmentation, 2020.

[YPR+19] Greg Yang, Jeffrey Pennington, Vinay Rao, Jascha Sohl-Dickstein, and Samuel S. Schoenholz. A mean field theory of batch normalization. In *International Conference on Learning Representations*, 2019.

[ZCDLP18] Hongyi Zhang, Moustapha Cisse, Yann N. Dauphin, and David Lopez-Paz. mixup: Beyond empirical risk minimization, 2018.

[ZDM19] Hongyi Zhang, Yann N. Dauphin, and Tengyu Ma. Fixup initialization: Residual learning without normalization, 2019.