High Performance Modeling, Knowledge Distillation

* [Distributed training](https://www.tensorflow.org/guide/distributed_training)
* [Data parallelism](https://arxiv.org/abs/1806.03377)
* [Pipeline parallelism](https://ai.googleblog.com/2019/03/introducing-gpipe-open-source-library.html)
* [GPipe](https://arxiv.org/abs/1811.06965)
* [Knowledge distillation](https://arxiv.org/pdf/1503.02531.pdf)
* [Q&A case study](https://arxiv.org/pdf/1910.08381.pdf)