NeurIPS Summary

CMS MLJC, 10th Dec 2018

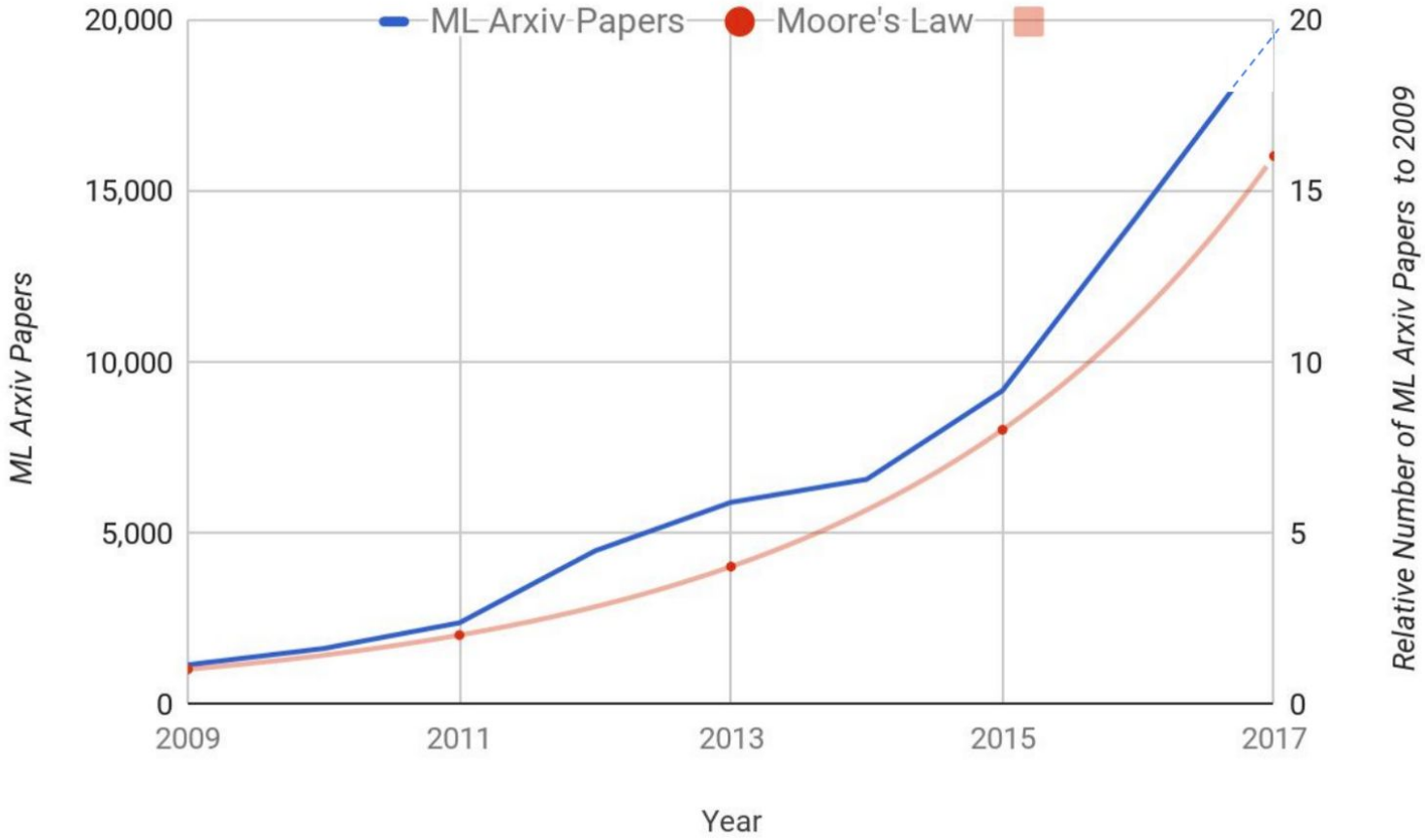
General Comments

- In Montreal, Canada
- Officially changed name NIPS -> NeurIPS
- Number of Submissions: > 4000
- ~ 800 accepted
- Sold out in < 12 minutes

Plenary Talks: Highlights

- Unsupervised Learning by Alex Graves: https://www.youtube.com/watch?v=rjZCjosEFpI
- Visualization in ML: UMAP:
  - Faster than TSNE
  - Seems to capture global structure better
  - Clusters are pulled away more efficiently

ML Arxiv Papers per Year



Best contributions according to reviewers

- Non-Delusional Q-learning and Value Iteration - Lu et. al.
- Optimal Algorithms for Non-Smooth Distributed Optimization in Networks - Scaman et.al.
- Nearly Tight Sample Complexity Bounds for Learning Figures of Gaussians via Sample Compression Schemes - Ashtiani et. al.
- **Neural Ordinary Differential Equations - Chen et. al.**

General Contributions to Bookmark

- “Recent Advances in Autoencoder-Based Representation Learning”: https://www.nari.ee.ethz.ch/commth/pubs/files/autoenc2018.pdf
- “Understanding Batch Normalization”: https://arxiv.org/abs/1806.02375
- “The Effect of Network Width on the Performance of Large-batch Training”: https://arxiv.org/abs/1806.03791
- “How to Start Training: The Effect of Initialization and Architecture”: https://arxiv.org/abs/1803.01719
- “L4: Practical loss-based stepsize adaptation for deep learning”: https://arxiv.org/abs/1802.05074
- ...

Isolation Sources of Disentanglement of VAEs

- Disentangled representation study

$$\mathcal{L}_\beta = \frac{1}{N} \sum_{n=1}^{N} \left(\mathbb{E}_q [\log p(x_n|z)] - \beta \, \mathbf{KL} \left(q(z|x_n) || p(z) \right) \right)$$

Isolation Sources of Disentanglement of VAEs

- ELBO decomposed to show magnify loss on total correlation between latent variables

$$\mathbb{E}_{p(n)}\Big[\mathrm{KL}\big(q(z|n)||p(z)\big)\Big] = \underbrace{\mathrm{KL}\left(q(z,n)||q(z)p(n)\right)}_{\text{(i) Index-Code MI}} + \underbrace{\mathrm{KL}\big(q(z)||\prod_j q(z_j)\big)}_{\text{(ii) Total Correlation}} + \underbrace{\sum_j \mathrm{KL}\left(q(z_j)||p(z_j)\right)}_{\text{(iii) Dimension-wise KL}} \tag{2}$$

$$\mathcal{L}_{\beta-\mathrm{TC}} := \mathbb{E}_{q(z|n)p(n)}[\log p(n|z)] - \alpha I_q(z;n) - \beta\ \mathrm{KL}\big(q(z)||\prod_j q(z_j)\big) - \gamma \sum_j \mathrm{KL}\left(q(z_j)||p(z_j)\right) \tag{4}$$