The College Of
WILLIAM & MARY

“Literature” Review

Alexander Radovic
College of William and Mary

Where to start?

You don't need a formal education in ML to use its tools. But it doesn't hurt to work through a online textbook or course. Here are a few I think would be fun & useful:

- The Coursera ML Course a very approachable introduction to ML, walks you through implementing core tools like backpropagation yourself
- CS231n: Convolutional Neural Networks for Visual Recognition another stanford course focused on NNs for "images", a great place to start picking up practical wisdom for our main use case
- Deep Learning With Python a book from the creator of keras, a great choice if you're planning to primarily work in python

Where do I get my news?

Twitter, slack, and podcasts are the only way I've found to navigate the vast amount of ML literature out there.

hardmaru Retweeted

Thomas Lahore @evolvingstuff · 10h

Wish more papers focused on understanding than results

Understanding Hidden Memories of Recurrent Neural Networks

arxiv.org/abs/1710.10777

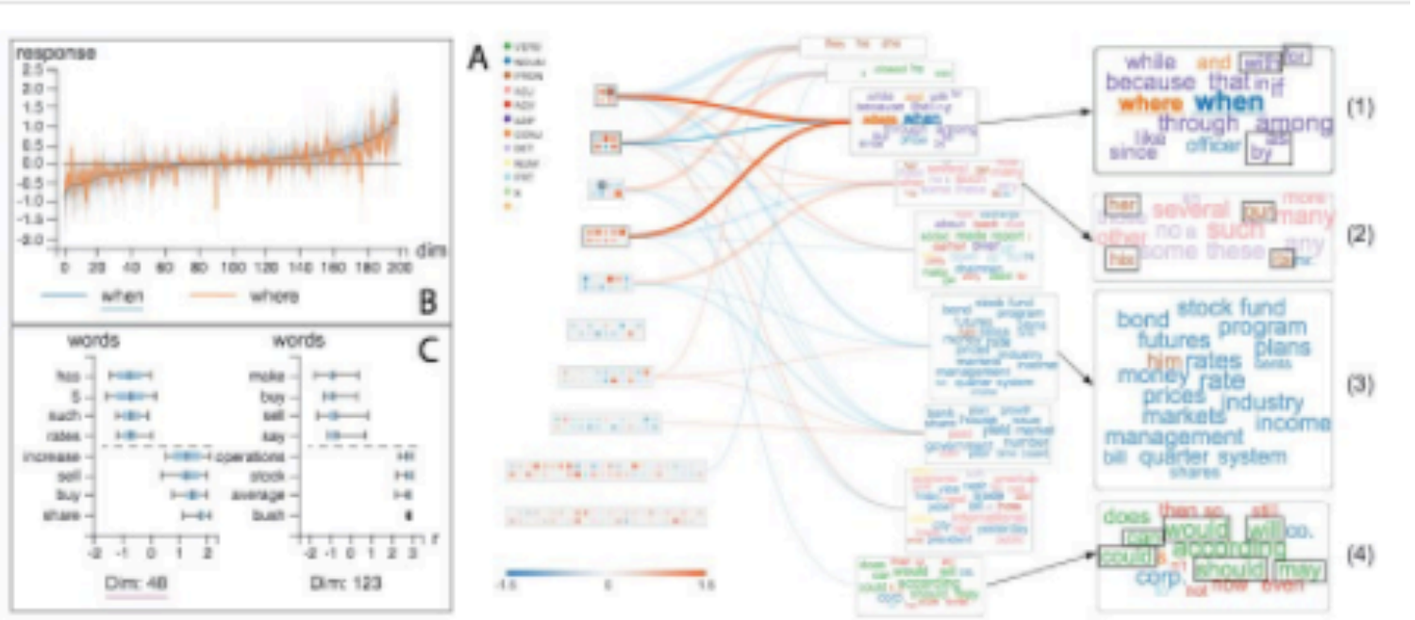


Figure 8: Co-cluster visualization of the last layer of LSTM-Small. (A): The co-Clustering layout with words colored according to POS tags. (B): The detail view showing model's responses to "when" and "where" in overlay manner. (C): Words interpretations of the 48th and 123rd cell states.

1 85 288

Kyle Cranmer
@KyleCranmer

Following

"Quantum machine goes in search of the Higgs boson" with quotes from yours truly @NatureNews

Quantum machine goes in search of the Higgs boson

D-Wave system shows quantum computers can learn to detect particle signatures in mountains of data, but doesn't outpace conventional methods — yet.

nature.com

1:51 PM - 19 Oct 2017

1 Retweet 7 Likes

2 1 7

Tweet your reply

Daniel Whiteson @DanielWhiteson · Oct 19

Replying to @KyleCranmer @NatureNews

QM tries to solve an already solved problem, comes up short. Published in Nature!

1 1

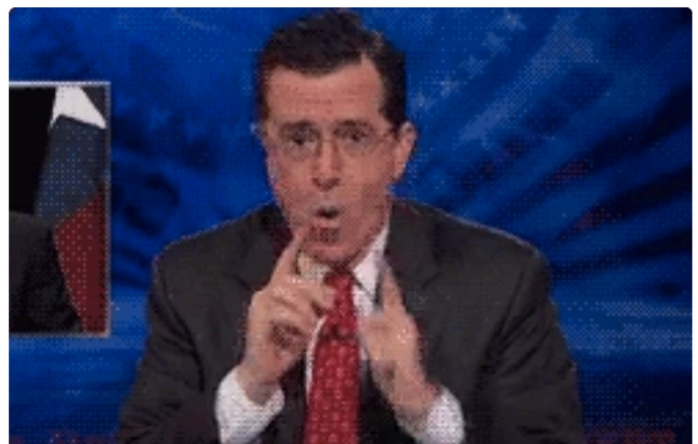
André David @DrAndreDavid · Oct 19

QC, not QM. Don't badmouth QM!

1

Michela Paganini @WonderMicky · Oct 20

Replying to @KyleCranmer @NatureNews



Where do I get my news?

Twitter, slack, and podcasts are the only way I've found to navigate the vast amount of ML literature out there.

hardmaru Retweeted

Thomas Lahore @evolvingstuff · 10h

Wish more papers focused on understanding than results

Understanding Hidden Memories of Recurrent Neural Networks

Daniel Whiteson @DanielWhiteson · Oct 19

Replying to @KyleCranmer @NatureNews

QM tries to solve an already solved problem, comes up short. Published in Nature!

Figure 8: Co-cluster visualization of the last layer of LSTM-Small. (A): The co-Clustering layout with words colored according to POS tags. (B): The detail view showing model's responses to "when" and "where" in overlay manner. (C): Words interpretations of the 48th and 123rd cell states.

1 85 288

Kyle Cranmer @KyleCranmer Following

"Quantum machine goes in search of the Higgs boson" with quotes from yours truly @NatureNews

Quantum machine goes in search of the Higgs boson

D-Wave system shows quantum computers can learn to detect particle signatures in mountains of data, but doesn't outpace conventional methods — yet.

nature.com

1:51 PM - 19 Oct 2017

1 Retweet 7 Likes

2 1 7

Tweet your reply

Daniel Whiteson @DanielWhiteson · Oct 19

Replying to @KyleCranmer @NatureNews

QM tries to solve an already solved problem, comes up short. Published in Nature!

1 1

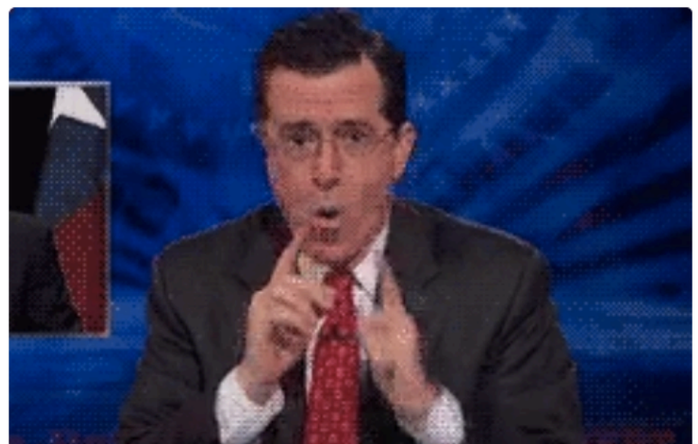
André David @DrAndreDavid · Oct 19

QC, not QM. Don't badmouth QM!

1

Michela Paganini @WonderMicky · Oct 20

Replying to @KyleCranmer @NatureNews



Where do I get my news?

Twitter, slack, and podcasts are the only way I've found to navigate the vast amount of ML literature out there.

hardmaru Retweeted

Thomas Lahore @evolvingstuff · 10h

Wish more papers focused on understanding than results

Understanding Hidden Memories of Recurrent Neural Networks
arxiv.org/abs/1710.10777

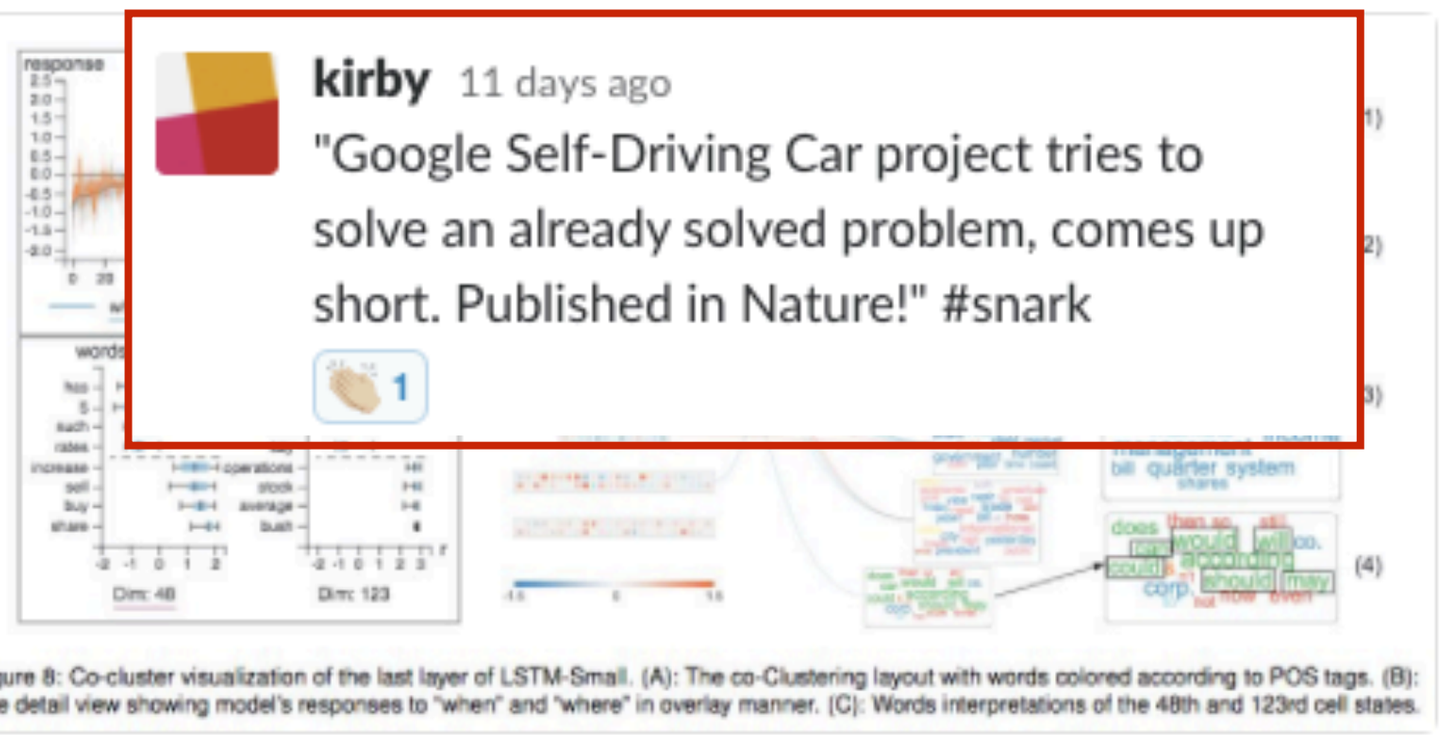


Figure 8: Co-cluster visualization of the last layer of LSTM-Small. (A): The co-Clustering layout with words colored according to POS tags. (B): The detail view showing model's responses to "when" and "where" in overlay manner. (C): Words interpretations of the 48th and 123rd cell states.

1 85 288

Thread
kirby, perdue, and you

alexanderradovic Oct 20th at 10:49 AM
in #journalclub

https://twitter.com/KyleCranmer/status/921086294448537600 the shade thrown in the replies to this tweet amused me

> **Kyle Cranmer** @KyleCranmer
> "Quantum machine goes in search of the Higgs boson" with quotes from yours truly @NatureNews http://www.nature.com/news/quantum-machine-goes-in-search-of-the-higgs-boson-1.22860
> Twitter | Oct 19th at 1:51 PM

2 replies

kirby 11 days ago
"Google Self-Driving Car project tries to solve an already solved problem, comes up short. Published in Nature!" #snark
1

perdue 11 days ago
LOL

Reply...

Also send to **#journalclub** Send

Where do I get my news?

Specifically I would recommend:

- Joining the fermilab machine learning slack
- Listening to Talking Machines podcast
- Following some great people on twitter:
  - Hardmaru @hardmaru, google brain resident, active & amusing with a focus on generative network work
  - Francois Chollet @fchollet, google based keras author, sometimes has interesting original work
  - Andrej Karpathy @karpathy, tesla director of ai, co-founder of first DL course at stanford
  - Kyle Cranmer @KyleCranmer, ATLAS NYU professor, helping lead the charge on DL in the collider would with lots of excellent short author papers
  - Gilles Loupe @glouppe, ML Associate Professor at the Université de Liège, a visiting scientist at CERN and often co-author with Kyle

Fun “Physics” Paper

So what should you read from recent HEP ML work?
https://arxiv.org/abs/1402.4735 the Nature paper that showed in MC that DNNs could be great for physics analysis
https://arxiv.org/abs/1604.01444 first CNN used for a physics result, should be familiar!

Can we train with less bias?
https://arxiv.org/abs/1611.01046 uses an adversarial network
https://arxiv.org/pdf/1305.7248.pdf more directly tweaking loss functions

RNNs for b-tagging and jet physics:
https://arxiv.org/pdf/1607.08633 first look at using RNNs with Jets
https://arxiv.org/abs/1702.00748 using recursive and recurrent neural nets for jet physics
ATLAS Technote first public LHC note showing they are looking at really using RNNs for b-tagging, CMS close behind

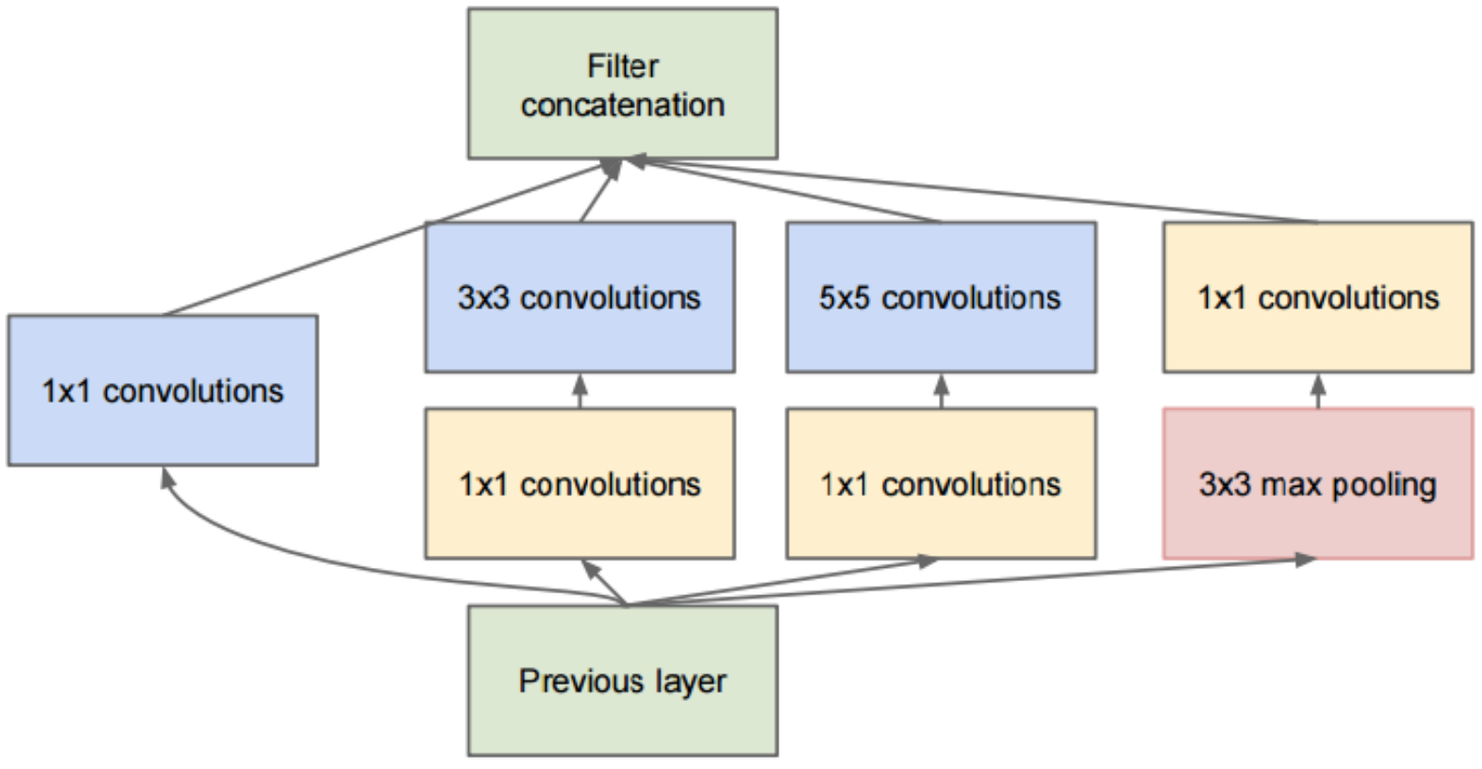
GANs for fast MC:
https://arxiv.org/abs/1705.02355 PoC for EM showers in calorimeters

CNN Papers

Our CNN for ID network is still very much inspired by the first googlenet:

https://arxiv.org/pdf/1409.4842v1.pdf

which introduces a specific network in network structure called an inception module which we've found to be very powerful.



CNN Papers

Our CNN for ID network is still very much inspired by the first googlenet:

https://arxiv.org/pdf/1409.4842v1.pdf

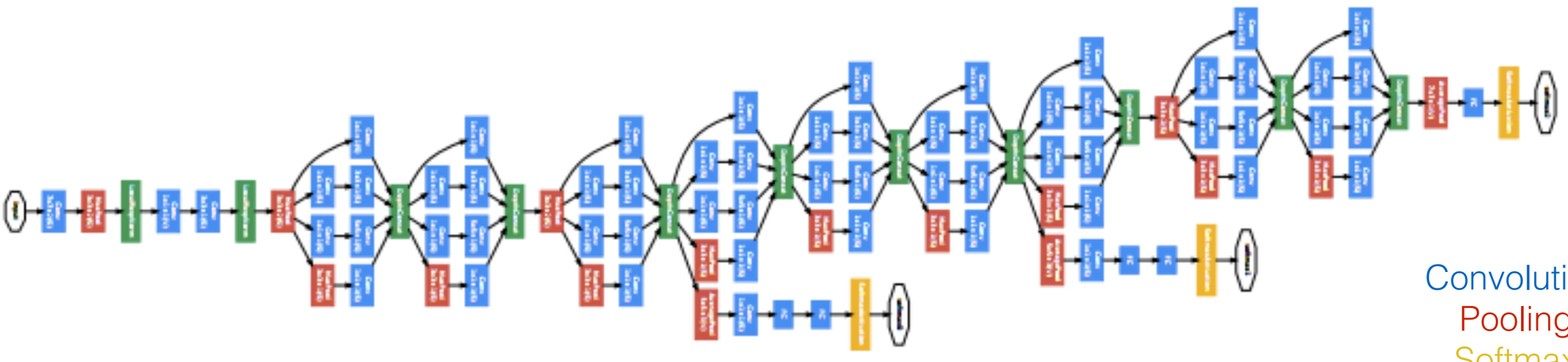
which introduces a specific network in network structure called an inception module which we've found to be very powerful.

CNN Papers

Our CNN for ID network is still very much inspired by the first googlenet:

https://arxiv.org/pdf/1409.4842v1.pdf

which introduces a specific network in network structure called an inception module which we've found to be very powerful.



The “GoogleNet” circa 2014

CNN Papers

Related to that paper are a number of papers charting the rise of the "network in network model", and advances in the googlenet that we've started to explore:
https://arxiv.org/abs/1312.4400 introduces the idea of networks in networks
http://arxiv.org/abs/1502.03167 introduces batch normalization which speeds training
http://arxiv.org/pdf/1512.00567.pdf smarter kernel sizes for GPU efficiency
http://arxiv.org/abs/1602.07261 introducing residual layers which enables even deeper networks

CNN Papers

We've also started to play with alternatives to inception modules inspired by some recent interesting models:

- https://arxiv.org/abs/1608.06993 the densenet which takes the idea of residual connections to an extreme conclusion
- https://arxiv.org/pdf/1610.02357.pdf replacing regular convolutions with depthwise separable ones under the hypothesis that 1x1 convolutional operations power the success of the inception module

CNN Papers

Or changing core components like the way we input an image or the activation functions we use

- https://arxiv.org/pdf/1706.02515.pdf an activation seems to work better than batch normalization for regularizing weights
- https://arxiv.org/abs/1406.4729 can we move to flexible sized inputs images?

Image Segmentation Papers

Can we break our events down to components and ID them?

- https://arxiv.org/pdf/1411.4038 first of a wave of cnn powered pixel-by-pixel IDS
- https://arxiv.org/abs/1505.04597 an example of where the task has been reinterpreted as an encoder/decoder task, with some insight from residual connection work, has worked very well for uboone
- https://arxiv.org/pdf/1611.07709.pdf part of work to ID objects in an image rather than individual pixels