reddit MACHINELEARNING   **comments**   rela Want to join? Log in or sign up in seconds. | **English**

▲
451
▼

! ## AMA Andrew Ng and Adam Coates (self.MachineLearning)

submitted 5 months ago by **andrewyng**

Dr. Andrew Ng is Chief Scientist at Baidu. He leads Baidu Research, which includes the Silicon Valley AI Lab, the Institute of Deep Learning and the Big Data Lab. The organization brings together global research talent to work on fundamental technologies in areas such as image recognition and image-based search, speech recognition, and semantic intelligence. In addition to his role at Baidu, Dr. Ng is a faculty member in Stanford University's Computer Science Department, and Chairman of Coursera, an online education platform (MOOC) that he co-founded. Dr. Ng holds degrees from Carnegie Mellon University, MIT and the University of California, Berkeley.

Dr. Adam Coates is Director of Baidu Research's Silicon Valley AI Lab. He received his PhD in 2012 from Stanford University and subsequently was a post-doctoral researcher at Stanford. His thesis work investigated issues in the development of deep learning methods, particularly the success of large neural networks trained from large datasets. He also led the development of large scale deep learning methods using

search [🔍]

this post was submitted on 13 Apr 2015
451 points (98% upvoted)

shortlink: `https://redd.it/32ihpe`

| username ✳ | password ✳ |

☐ remember me   reset password   [login]

Submit a new link ›

Submit a new text post ›

MachineLearning

[subscribe] 45,655 readers

○ 108 users here now

News, Research Papers, Videos, Lectures, Softwares and Discussions on:

- Machine Learning
- Data Mining
- Information Retrieval
- Predictive Statistics
- Learning Theory
- Search Engines
- Pattern Recognition
- Analytics

Journal Club

Current Discussion Thread
Current Voting Thread

Current Simple Question Thread

AMAs:

Andrew Ng and Adam Coates (4/15/2015)
Jürgen Schmidhuber (3/4/2015)
Geoffrey Hinton (11/10/2014)
Michael Jordan (9/10/2014)
Yann LeCun (5/15/2014)
Yoshua Bengio (2/27/2014)

Beginners:

distributed clusters and GPUs. At
Stanford, his team trained
artificial neural networks with
billions of connections using
techniques for high performance
computing systems.

Please have a look at our FAQ and Link-Collection
Metacademy is a great resource which compiles
lesson plans on popular machine learning topics.
You could also try /r/mlclass or /r/MLQuestions

276 comments   share

top 200 comments show all 276

sorted by: **best** ▼

Related Subreddit :

- **Statistics**
- **Computer Vision**
- **Compressive Sensing**
- **NLP**
- **Jobs**
- **/r/datacleaning**

created by kunjaan   a community for 6 years

MODERATORS   message the moderators

kunjaan
kanak
cavedave naive
olaf_nij
BeatLeJuce

about moderation team »

[−] **RileyNat** 89 points 5 months ago

I am a big fan of your work Dr. Ng, your
coursera course was what introduced me to
Machine Learning. My question is do you think a
PhD or Masters degree is a strong requirement
for those who wish to do ML research in industry
or can a Bachelors and independent learning be
enough? Thanks.

permalink

   [−] **andrewyng** [S] 58 points 5 months ago

   Thank you RileyNat for taking the Coursera MOOC.

   Regarding the need for a degree in ML: Absolutely not! I think a PhD is one great way
   to learn about machine learning. But many top machine learning researchers do not
   have a PhD.

   Given my (Andrew's) background in education and in Coursera, I believe a lot in
   employee development. Thus at most of the teams I've led (at Baidu, and previously
   when I was leading Google's Deep Learning team/Google Brain) I invested a lot in
   training people to become expert in machine learning. I think that some of these
   organizations can be extremely good at training people to become great at machine
   learning.

   I think independent learning through Coursera is a great step. Many other software
   skills that you may already have are also highly relevant to ML research. I'd encourage
   you to keep taking MOOCs and using free online resources (like
   deeplearning.stanford.edu/tutorial). With sufficient self-study, that can be enough to
   get you a great position at a machine learning group in industry, which would then help
   further accelerate your learning.

   permalink   parent

      [−] **YashN** 18 points 5 months ago

      Best course i have ever taken. Felt we were part of something special, unique.

      permalink   parent

         [−] **ginger_beer_m** 7 points 4 months ago

         I was part of the early batch of the ML class before it was called Coursera.
         Andrew was sort of almost-crying in the last video, and it was a great journey

together.

permalink parent

load more comments (2 replies)

load more comments (1 reply)

[−] **hachidan05** 9 points 5 months ago

This question is very common. I would try to help answer this question. It depends what company you are trying to work for. The bigger the company the more they want to see a Masters or PhD degree although that is not always the case. The smaller/newer companies are more willing to accept Bachelors with independent learning. I was looking for entry jobs in the data science field one day and notice a company write out in the job description "Online degree or certificate can replace 1 year of relevant job experience." Just think about this, there are more and more companies wanting people with data analytic/mining skills to get an edge in their respective industries. I believe online learning is also on the rise. Sooner or later companies will accept online certificate (free or not). Masters/PhD are for people wanting to focusing on a area within ML. Bachelors and independent learning is for people who want to get their feet in the door then get the small company pay for your Masters/PhD. Some smaller company pay competitively and raises are based on performance of the company/your contribution.

permalink parent

[−] **andrewyng** [S] 31 points 5 months ago

I agree that the newer companies---ones that know how to evaluate machine learning talent---care more about your ability, and less about the credential (such as MS or PhD). For example, at Baidu Research, we do hire top machine learning researchers and machine learning engineers that don't have a graduate degree, but have great software skills and have knowledge of ML from elsewhere.

Over time, companies are also increasingly valuing certificates earned from MOOCs.

permalink parent

[−] **Refefer** 2 points 5 months ago

There really isn't much specialization within ML at the Masters level, even at the Top 10 schools - that's mostly reserved for PhDs.

permalink parent

load more comments (3 replies)

[−] **InsideAndOut** 6 points 5 months ago

If you want to do ML research, where research means developing new algorithms, methods, anything not in the already present "ML cookbooks" - you need a PhD (or a Master's and significant experience in research).

From what I've heard from a few friends currently working in Facebook & Google - there they don't let you touch research (or even ML) without a PhD in the field.

From my personal experience, I got a few job opportunities (NLP, ML), and on each of the interviews, the company / team leader already had a set plan for which methods will be used for the problem.

The only options I see possible, as /u/hachidan05 already wrote, is aiming for

startups/smaller companies, which usually have lower standards - and finding one that will allow you to do research for them.

Alternatively, set up a strong GitHub account. Employers often check those things, and if you have coded SVM's, regression models, clustering from scratch (and applied it successfully to some known datasets), that could be proof enough of your skill.

My background - MSc in Computer Science - ML+NLP & 1 year of work in the field

Since I'm kind of late to the party, I'll piggyback off of your comment and try to ask prof. Ng a question as well -

1. Which European universities do you consider best for ML? Or, more specific, which professors do you consider the best in Europe?

My "field of expertise" is ML + NLP with a bit of information retrieval.

I'm planning to apply for a PhD and this would be significant help - thank you.

permalink   parent

load more comments (1 reply)

[−] **gardinal** 5 points 5 months ago

Same question. Very relevant because a huge number of students are taking your ML courses to start with.

permalink   parent

[−] **hadrielle** 2 points 5 months ago

Thanks for asking this question. I'm in exactly same situation! But was afraid to ask it.

permalink   parent

[−] **[deleted]** 66 points 5 months ago

What motivates some big companies to publish their ML tricks, like e.g. the recent Batch Normalization from Google? Aren't they giving away their secret sauce to competitors?

Do you think the published results are just the tip of the iceberg, and the very best findings are kept secret?

permalink

[−] **andrewyng** [S] 44 points 5 months ago

As a research organization, Baidu Research and others want to be part of the community, and we want to learn from as well as contribute to it. Of course, publishing also helps us attract talent, and also give our team better internal and external visibility. But underlying this is that we're researchers and just want to invent ideas that help make the world a better place!

Having said that, the mission of the Baidu's AI Lab is to develop hard AI technologies that let us impact hundreds of millions of users. Thus our focus is on developing and shipping technologies. It's just that we're pretty open and transparent and are happy to publish a lot of what we learn along the way.

(By the way, Adam Coates and I are sitting together, so you should assume all these answers are written by both of us.)

permalink   parent

load more comments (1 reply)

[−] **londons_explorer** 13 points 5 months ago

To hire the best researchers they have to demonstrate how world class their research is, which in turn requires publishing lots of good papers.

Google publish papers about the majority of advancements in ML. The thing they rarely talk about is which specific services within Google use ML. For example, there are no papers about machine learning in web search.

permalink parent

> [−] **barmaley_exe** 18 points 5 months ago
>
> AFAIK, Google doesn't (or tends to not) use a lot of machine learning for search. See this question on Quora: http://www.quora.com/Why-is-machine-learning-used-heavily-for-Googles-ad-ranking-and-less-for-their-search-ranking
>
> permalink parent
>
> > load more comments (2 replies)

[−] **Smallpaul** 2 points 5 months ago

I wonder whether these techniques are actually patented so that Google profits if others build upon them (because they can demand licensing fees).

permalink parent

> [−] **alexmlamb** 3 points 5 months ago
>
> To my knowledge there aren't any credible patents in deep learning.
>
> This is unlike much of computer vision, which has a minefield of patents holding back progress.
>
> permalink parent
>
> > [−] **Smallpaul** 2 points 5 months ago
> >
> > Patents take a year or two to go through the system.
> >
> > permalink parent
> >
> > > load more comments (3 replies)
>
> load more comments (2 replies)

load more comments (2 replies)

[−] **pmill10** 37 points 5 months ago*

In a couple prior AMA's, Juergen Schidhuber and Michael Jordan were both asked to suggest reading lists for students beginning graduate school. I was surprised by how little their lists overlapped. I didn't expect them to recommend the same books, but at least books on the same topics.

Michael Jordan recommended books on more pure math/stats topics (Bayesian statistics, frequentists statistics, mathematical statistics, functional analysis, measure theory). Schmidhuber recommended books on information theory, theory of computation, algorithmic information theory.

Which of these topics would you recommend focusing on, and why? What books do you recommend to students entering your lab at Stanford?

https://www.reddit.com/r/MachineLearning/comments/2xcyrl/i_am_j%C3%BCrgen_schmidhuber_ama/coz4w3o

http://www.reddit.com/r/MachineLearning/comments/2fxi6v/ama_michael_i_jordan/ckdqzp
h

permalink

[−] **doodhwaala** 65 points 5 months ago

@andrewyng

What kind of self projects and follow up courses would you recommend after the Coursera
ML course?

permalink

    [−] **andrewyng** [S] 54 points 5 months ago

    Here're a few common paths: 1. Many people are applying ML to projects by themselves
    at home, or in their companies. This helps both with your learning, as well as helps
    build up a portfolio of ML projects in your resume (if that is your goal). If you're not
    sure what projects to work on, Kaggle competitions can be a great way to start. Though
    if you have your own ideas I'd encourage you to pursue those as well. If you're looking
    for ideas, check out also the machine learning projects my Stanford class did last year:
    http://cs229.stanford.edu/projects2014.html I'm always blown away by the creativity
    and diversity of the students' ideas. I hope this also helps inspire ideas in others! 2. If
    you're interested in a career in data science, many people go on from the machine
    learning MOOC to take the Data Science specialization. Many students are successfully
    using this combination to start off data science careers.
    https://www.coursera.org/specialization/jhudatascience/1

    permalink   parent

      load more comments (1 reply)

    [−] **DreamFlasher** 7 points 5 months ago

    There seems to be some "general ML wisdom" which is not taught in courses like yours
    or Daphne Kollers PGM, but it enables people (experts) in the field to understand each
    other's research/presentations. How/where can one acquire this knowledge?

    permalink   parent

[−] **iwantedthisusername** 66 points 5 months ago

Hinton seems to think that the next neural abstraction after the layer is the artificial cortical
column. Have you done any work toward this end goal?

Also what are your thoughts on HTM and the CLA (Numenta)

permalink

    [−] **iwantedthisusername** 7 points 5 months ago

    Why did they ignore this question? It seems like plenty of others wanted to hear the
    answer.

    permalink   parent

      load more comments (1 reply)

    [−] **tabacof** 2 points 5 months ago

    He is talking about Hinton's work on capsules:
    http://www.reddit.com/r/MachineLearning/comments/2lmo0l/ama_geoffrey_hinton/clyj
    6mt

There is a 2011 article about it but it's probably outdated:
http://www.cs.toronto.edu/~fritz/absps/transauto6.pdf

permalink   parent

[−] **NicolasGuacamole**   2 points 5 months ago

Do you have any links to Geoff speaking on this?

I'd like to understand it before the answering starts

permalink   parent

[−] **blindConjecture**   4 points 5 months ago*

That sounds a lot more like Jeff Hawkins than Geoff Hinton. Are you sure you're not getting them mixed up?

Edit: It looks like you did actually mean Hinton after all. Thanks /u/tabacof for clearing up the confusion.

permalink   parent

load more comments (2 replies)

> discussions in /r/MachineLearning   X
>
> · 1 comment
>
> 3 minute summary of The Empire Strikes Back (applying automatic text summarization to film)

[−] [deleted] 26 points 5 months ago

What do you see as the road map for ML in the next decade?

permalink

[−] [deleted] 22 points 5 months ago

Your much-cited 2011 AISTATS paper showed k-means with ZCA whitening to be competitive or superior to other, more complex, unsupervised natural image feature learning approaches.

Since then, denoising AEs, marginalized denoising AEs and other models appeared, as well as better ways to optimize deep nets, although I haven't seen an updated study like yours. Would you still expect k-means to be competitive in this domain?

permalink

[−] **adamcoates**   Director of Baidu Research   9 points 5 months ago

I think part of the value in the K-means approach was its simplicity and ability to scale up well. How K-means compares to current unsupervised learning methods isn't clear to me, but the lasting insight from that work has been the importance of scalability. Even though K-means is very simple, you could often make it competitive by building very large models.

In supervised deep learning, many of the algorithms that we use are still very simple (e.g., backpropagation), yet by scaling them up we can often outperform more sophisticated methods. In the AI Lab, we have a lot of great systems researchers (e.g., Bryan Catanzaro, who created CuDNN) that work on scaling up deep learning algorithms, etc. based on this insight.

permalink   parent

[−] **soulslicer0**   1 point 5 months ago

In this paper, hard labelling versions K-means and GMM are used. Are there any advantages to user soft-labelling, where we assign different weights to each cluster?

permalink   parent

[–] **shaggorama** 1 point 5 months ago

they do use a kind of soft labeling k-means in that paper, and it delivered the best performance of all methods tested. Also, the GMM approach they used was soft labeling: they didn't do a hard labeling GMM.

permalink   parent

[–] **eldeemon** 29 points 5 months ago

Hi Andrew and Adam! Many thanks for taking the time for this!

(1) What are your thoughts on the role that theory is to play in the future of ML, particularly as models grow in complexity? It often seems like that the gap between theory and practice is widening.

(2) What are your thoughts on the future of unsupervised learning, especially now that (properly initialized and regularized) supervised techniques are leading the pack? Will layer-by-layer pretraining end up as a historical footnote?

permalink

[–] **andrewyng** [S] 23 points 5 months ago

Hi Eldeemon,

Great question. I think that 50 years ago, CS theory was really driving progress in CS practice. For example, the theoretical work figuring out that sorting is O(n log n), and Don Knuth's early books, really helped advance the field. Today, there're some areas of theory that're still driving practice, such as computer security: If you find a flaw in crypto and publish a theoretical paper about it, this can cause code to be written all around the world.

But in machine learning, progress is increasingly driven by empirical work rather than theory. Both still remain important (for example, I'm inspired by a lot of Yoshua Bengio's theoretical work), but in the future I hope we can do a better job connecting theory and practice.

As for unsupervised learning, I remain optimistic about it, but just have no idea what the right algorithm is. I think layer-by-layer pretraining was a good first attempt. But it really remains to be seen if researchers come up with something dramatically different in the coming years! (I'm seeing some early signs of this.)

permalink   parent

[–] **sixwings** 6 points 5 months ago

> As for unsupervised learning, I remain optimistic about it, but just have no idea what the right algorithm is. I think layer-by-layer pretraining was a good first attempt. But it really remains to be seen if researchers come up with something dramatically different in the coming years! (I'm seeing some early signs of this.)

Can you share those early signs with the rest of us?

permalink   parent

[–] **willwill100** 17 points 5 months ago

What do either of you think the current big bottlenecks in AI are that are preventing the next big leap forward?

permalink

[−] **wearing_theinsideout** 8 points 5 months ago

Hey Andrew, huge fan of your work, mainly Machine Learning Coursera course that basically started my interest in ML area.

Question: I have seen that your work is focused in DL, however I have not seen or read any work of yours focusing on Recurrent Neural Networks (RNN). Works in this area like the one that has been done by Schmidhuber with Long Short-Term Memories (LSTM) are very famous and started to win some contests. Have you never thought about working and researching with RNNs? With your experience, can you point some pros and cons of RNNs?

Thanks a lot!

permalink

> [−] **andrewyng** [S] 10 points 5 months ago
>
> I think RNNs are an exciting class of models for temporal data! In fact, our recent breakthrough in speech recognition used bi-directional RNNs. See http://bit.ly/deepspeech We also considered LSTMs. For our particular application, we found that the simplicity of RNNs (compared to LSTMs) allowed us to scale up to larger models, and thus we were able to get RNNs to perform better. But at Baidu we are also applying LSTMs to a few problems were there is are longer-range dependencies in the temporal data.
>
> permalink   parent

[−] **OnlySpeaksLies** 18 points 5 months ago

What are your thoughts on ML competitions (the most well-known example being kaggle)? And more generally, do you think gamification is beneficial to (ML) research?

permalink

load more comments (8 replies)

[−] **idontuseredditalot** 24 points 5 months ago*

Can you touch on what makes Baidu, in terms of the work being done (applications of the research etc.) and the work environment (autonomy, benefits, career growth / mobility etc.), a more attractive workplace for machine learning and AI engineers vs. other companies like Google or Facebook?

permalink

> [−] **andrewyng** [S] 15 points 5 months ago
>
> I think Baidu, Google and Facebook are all great places to work!
>
> I don't want to compare Baidu against any other company (since I think they're all great). But Baidu Research is very much a startup environment. With ~40 people in our Silicon Valley team, we tend to act with the nimbleness of a startup of a commensurate size (albeit with the access to computational power and data of a $75B company). We also invest a lot in employee development, and so I see that people here are all working hard and learning rapidly about deep learning, HPC, etc. I think these things make the best possible combination for driving machine learning research, which is why both of us (Adam & Andrew) had decided to join Baidu.
>
> permalink   parent

[−] **amitm02** 7 points 5 months ago

Hi, big fan here :)

- Is there a unified "machine learning algorithms theory" that connects the different methods (e.g. logistic regression, SVMs, boosting, random forests, deep networks etc')?
- Are the capacities/coverage/descriptive power of the different methods nested (one system always have a more descriptive power than the other) or disjointed/overlap (one system can describe really well one pattern while the other does better for a different one)?
- Can we show how each method relates to another by building an analogous system + adding or removing constraints on that system?

Thanks in advance, Amit

permalink

> [−] **letitgo12345**   3 points 5 months ago
>
> For (1), a good tutorial, I've found is this http://yann.lecun.com/exdb/publis/pdf/lecun-06.pdf
>
> permalink   parent

[−] **huitseeker**   6 points 5 months ago

In your GTC15 keynote (around the 13:00 mark), you relay Jensen's message "We've done a good job of convincing people to use single GPUs, but [..] this isn't about buying a single GPU any more, it's about buying 32 GPUs".

Nonetheless, a lot of results in **distributed** deep network training are proprietary, from Google's in 2012 to Netflix's in 2014.

- Would you have a few *accessible* pointers on where to go look for more information on the **distributed** training of deep networks ?
- Do you think open-sourcing of projects on this particular subject is relatively close on the horizon ?

permalink

[−] **shaggorama**   21 points 5 months ago

I find that I -- and most other data scientists I know -- generally have one or several analytics/programming hobby projects unrelated to work that we are tinkering with. What are some bench/hobby projects either or both of you have been tinkering with lately (assuming that's even a thing you do)?

permalink

[−] **valexiev**   20 points 5 months ago

Hello, Dr. Ng and Dr. Coates! First off, thank you for taking the time to answer our questions!

1. My first question is about career advice. My goal is to do Machine Learning research in industry. However, due to a variety of circumstances, I was unable to go to grad school full time, though I'm currently studying for my Online Masters at Georgia Tech part-time while working as a software developer. One of my biggest regrets about this setup is that I'm not able to do research or collaborate with a professor, which I think is very important for my future career. How can I overcome this drawback? In your experience, how is independent research seen in industry and academia? What would be some ways in which I can get in contact and possibly collaborate with the research

community?

2. I know that the recent AI Doomsday prophesies by Elon Musk, Stephen Hawking and others have been met with justified amusement and skepticism from the majority of AI and ML practitioners. In fact, I remember reading an article in which Dr. Ng outlined why he's not spending any time worrying about it because the current technology is very far from achieving something like that. In your opinion, what would be some achievements in the AI field that would signal to you that AGI is close to becoming a reality?

Again, thank you for taking the time for this AMA and, Dr. Ng, thank you for your excellent Coursera course on ML!

permalink

[−] **andrewyng** [S] 25 points 5 months ago

Hi Valexiev,

Thrilled to hear that you want to do machine learning research in industry! If you have a strong portfolio of projects done through independent research, this counts for a lot in industry. For example, at Baidu Research we hire machine learning researchers and machine learning engineers based only on their skills and abilities, rather than based on their degrees; and, past experience (such as demonstrated in a portfolio of projects) helps a lot in evaluating their skills.

I think that mastering the basics of machine learning (for example, through MOOCs, and free resources like deeplearning.stanford.edu/tutorial) would be the best first step. After that, I'd encourage you to find projects either in your company or by yourself to work on and to use this to keep learning as well as to build up your portfolio. If you don't know where to start, Kaggle is a reasonable starting place; though eventually I'd encourage you to also identify and work on your own projects. In the meantime, do keep reaching out to professors, and attend local meetups, and try to find a community.

This is often enough to find you a position to do machine learning work in a company, which then further accelerates your learning.

permalink parent

[−] **valexiev** 1 point 5 months ago

Thank you for your response!

permalink parent

[−] **SunnyJapan** 13 points 5 months ago

What are the most exciting things that are happening in Deep Learning field now?

permalink

[−] **adagrads** 6 points 5 months ago

Does it bother you that Baidu censors search results for China? Or that their government has weaponized Baidu Analytics? See: https://en.greatfire.org/blog/2015/mar/chinese-authorities-compromise-millions-cyberattacks

I ask this because it seems like deep computer vision, for example, will significantly enhance their capabilities for censorship of free speech, among other things.

permalink

[−] **NicolasGuacamole** 14 points 5 months ago

What do you wish you had known at the start of your career?

permalink

[−] **andrewyng** [S] 14 points 5 months ago

One of the things both of us (Adam & Andrew) talk about frequently is the impact of research. At Baidu, our goal is to develop hard AI technologies that impact hundreds of millions of users. Over time, I think we've both learned to be more strategic, and to learn to see more steps out ahead--beyond just writing a paper--to plot a path to seeing our technology benefit huge numbers of people. These days, this is one of the things that really excite us about our work!

permalink parent

[−] **test3545** 13 points 5 months ago

Jürgen Schmidhuber QUOTE: *"Since BP was 3-5 decades old by then, and pattern deformations 2 decades, these results seemed to suggest that advances in exploiting modern computing hardware were more important than advances in algorithms."* [1]

Yann LeCun QUOTE: *"Basically we limited by computational power. So, the faster, you know, the next generation of Nvidia GPU will be the more progress we'll make."* [2]

What is your opinion about the matter?

[1] Juergen Schmidhuber, 2014, Deep Learning in Neural Networks: An Overview

[2] Yann LeCun, 2014, Convolutional Networks- Machine Learning for Computer Perception (Nvidia webinar, 2014)

permalink

[−] **andrewyng** [S] 11 points 5 months ago

I think the two key drivers of deep learning are: - Rise of computation. Not just GPUs, but now the migration toward HPC (high performance computing, aka supercomputers). - Rise of availability of data, because of the digitization of our society, in which increasing amounts of activity on computers/cellphones/etc. creates data.

Of course, algorithmic progress is important too, but this progress is enabled by the rise of computational resources and data.

I think though that the rise of computation isn't something we passively wait to let happen. In both of our (Adam+Andrew's) careers in deep learning, a lot of our success was because we actively invested to increase the computation available.

For example, in 2008, we built I think the first CUDA/GPU deep learning implementation, and helped lead the field to use GPUs. In 2011, I (Andrew) founded and led the Google Deep Learning team (then called Google Brain) to use Google's cloud to scale up deep learning; and this helped put it on industry's radar. In 2013, Adam, Bryan Catanzaro and others built the first HPC-style deep learning system, and this helped drive scaling another 1-2 orders of magnitude.

Finally, today at Baidu, we have a system's team that's developing what we think is the next generation of deep learning systems, using HPC techniques. If you're not familiar with HPC, it's a very different set of tools/people/conferences/methods than cloud computing, and this is giving us another big boost in computation. We think it's the

combination of HPC and large amounts of data that'll give us the next big increment in deep learning. For example, this is what enabled our recent breakthrough in speech recognition (http://bit.ly/deepspeech).

For more on the latest in deep learning+HPC, take a look at my (Andrew's) keynote at the GPU Technology Conference: http://www.ustream.tv/recorded/60113824

permalink parent

[−] **falconberger** 2 points 5 months ago

> In 2013, Adam, Bryan Catanzaro and others built the first HPC-style deep learning system, and this helped drive scaling another 1-2 orders of magnitude.

Any chance parts of this will get open-sourced?

permalink parent

[−] **AlcaDotS** 3 points 5 months ago

> [1] Juergen Schmidhuber, 2014, Deep Learning in Neural Networks: An Overview

For those interested, reference [1] points to section 5.18 (page 23)

permalink parent

[−] **llevar** 6 points 5 months ago

A common weakness of Coursera and edX MOOCs is that they are watered down superficial versions of live courses. Students are not asked to solve any hard problems for fear of losing the audience, but as a result are not able to really learn the content of the course in a way that will allow them to apply it in real life scenarios. There are very few exceptions like Daphne Koller's PGM course or the ML course from Caltech on edX.

Do you see any place for advanced Masters or PhD level courses on the Coursera platform, and if so, what steps are you taking to encourage their creation?

permalink

[−] **TMaster** 2 points 5 months ago

My experience is the opposite. I have much more interaction with the material on Coursera and understand it better than I do in offline universities because of Coursera's automation.

I see a very high correlation between making a course computer-based and my results in it. For me, barriers are lowered through the direct feedback of the quizzes and the automated direct assessment of your own code. Being able to selectively pause, rewind, and re-watch lectures has not been offered to me by offline universities and vastly improves understanding for me as well. When you're sitting in a hall with 100 students, I've found that fellow students *don't* like it when questions are asked, because everyone has different aspects they get stuck on, and what is unclear for you may well be clear to others. That's demotivating, but doesn't apply to Coursera.

In offline universities, it usually takes weeks for your results to get back to you, and by that time you've been put on new assignments already.

I don't know if MOOCs will overtake offline universities, but I do know that they are more effective for me.

permalink parent

[−] **llevar** 3 points 5 months ago

Thanks for weighing in. I don't think we actually disagree. I also like all the things you mentioned about MOOCs. My comment relates to the level of the material that is presented and the difficulty of the homeworks. I've completed close to two dozen Coursera and edX courses now and only a couple come anywhere near the level of complexity of higher level undergrad or graduate courses. This has mostly to do with the fact that a typical homework takes the shape of a multiple choice quiz that gives you 100 tries and can be completed in half an hour with only a vague understanding of the material. An upper level university course on the other hand involves independent problem solving and development of ideas - activities that require you to incorporate course concepts into your working memory.

I see the same Intro to Stats/Data Science or Single Variable Calculus popping up over and over again on Coursera but not a single Bayesian Inference, or Group Theory, or any other "Insert Advanced Subject Here". Having these would be quite nice as many on Coursera already have university degrees and are not well served by the innumerable introductory courses.

permalink   parent

[–] **TMaster** 2 points 5 months ago

Well, our experiences certainly don't match up, but note that I'm not necessarily arguing from an objective basis, just relaying my own experiences.

While there are ample courses that are less advanced on Coursera, the fact that they have the interaction I never got in university more than makes up for it for me. I'm fairly certain I simply learn more in any given Coursera course than offline course because of the reasons I've given.

> This has mostly to do with the fact that a typical homework takes the shape of a multiple choice quiz that gives you 100 tries and can be completed in half an hour with only a vague understanding of the material.

Try a course where you do need to do so, especially the courses where you're encouraged to code up solutions to problems. Try guessing a number then or getting to the answer with only partial understanding...

> I see the same Intro to Stats/Data Science or Single Variable Calculus popping up over and over again on Coursera but not a single Bayesian Inference, or Group Theory, or any other "Insert Advanced Subject Here".

I haven't even seen those (but I haven't looked since I'm supposed to be past that level anyway). Try cryptography. Algorithms. Ng's machine learning (did you do that one already?). Electrical engineering.

I do agree that more advanced courses would be better though.

permalink   parent

load more comments (1 reply)

[–] **wolajacy** 1 point 5 months ago

I agree, one can see a great gap between Andrew's course taught in Stanford and the Coursera version. A lot of math is ommited, theories are simplified - that's surely a big disadvantage of MOOC (i've taken a few other courses and all of them had that lack-of-deep-math problem).

permalink   parent

[−] **4rgento** 3 points 5 months ago

Dear Professor Ng,

> What are some problems that you think: "ML could be helpful here!" where you don'
> t think there is enough ML utilization?
>
> What was the most unexpected ML application that you saw?

Thanks for the time you spent on the ml-class at coursera.

permalink

[−] **elsonidoq** 6 points 5 months ago

Hi Dr. Ng! Thank you for your time. With all your experience, I would like to know how you handle model selection in deep convolutional architectures (eg, how many layers, kernel size, whether and where to use pooling, regularization, etc).

I am working on a ConvNet to classify text starting from word2vec embeddings, and even though I get decent results, I can not get to an overfitting architecture on the full dataset (it does overfit on a very very small sample) so I have the feeling I am actually underfitting. Any general tips on this?

Again, thank you very much!

Pablo

permalink

[−] **SuperFX** 12 points 5 months ago

Do you think neural networks will continue to be the dominant paradigm in ML, or will we see a swing back to greater diversity, with things like Bayesian nonparametrics and deep architectures constructed out of non-NN layers?

permalink

> [−] **llevar** 3 points 5 months ago
>
> Linear/logistic regression and k-means clustering are probably the dominant paradigms in ML, and likely will always be. There's just too much bang for the buck.
>
> permalink   parent

> [−] **FixDeineKabel** 4 points 5 months ago
>
> Neural networks are certainly not the dominant paradigm in ML.
>
> permalink   parent

>> [−] **ralphplzgo** 8 points 5 months ago
>>
>> aren't they in terms of state-of-the-art progress on numerous tasks?
>>
>> permalink   parent

>>> [−] **FixDeineKabel** 3 points 5 months ago
>>>
>>> They are state-of-the-art on a few supervised tasks for which there is a large amount of data. While that set of tasks is growing, machine learning is a large field and deep learning is still a relatively small part of it in both industry and academia.
>>>
>>> permalink   parent

[−] **OnlySpeaksLies** 2 points 5 months ago

Computer vision, reinforcement learning and language processing are very broad subjects though.. Of course, DL is often used in combination with some other method, but it is very prominent nowadays.

I'm not saying that you should just forget about Bayesian methods or manually constructed features, but, in my opinion, it is not unlikely that CNNs will be taught in CV courses soon.

permalink parent

[−] **FixDeineKabel** 2 points 5 months ago

Sure. Note that I did not say that neural networks were unimportant. I said that it is incorrect to call them the dominant paradigm in machine learning. I'd agree that CNNs will be taught in computer vision courses soon, and they obviously should be, but that doesn't change my point.

permalink parent

load more comments (1 reply)

load more comments (1 reply)

[−] **zdk** 3 points 5 months ago

Thank you! You'd never know it from this sub, since every other post is on some deep learning NN. In my field (computational biology) I never see enough data that would justify this approach.

permalink parent

[−] **NicolasGuacamole** 5 points 5 months ago

I'm an undergraduate with an extremely keen interest in Machine Learning and specifically Neural Networks; for someone young trying to get into the field what is the best piece of personal advice you could give them?

permalink

[−] **lamecicle** 8 points 5 months ago

Will you use/are you using ML to answer this AMA?

permalink

[−] **letoseldon** 6 points 5 months ago

Models of neural networks were originally inspired by biological systems but have since evolved beyond their original constraints.

Do you see any new findings from neuroscience inspiring future machine learning techniques, and is there any recent neuroscience research that you think is promising in terms of ML?

permalink

load more comments (1 reply)

[−] **kunjaan** 3 points 5 months ago

What has been the most surprising application of learning algorithms that you have seen?

permalink

[−] **letitgo12345** 11 points 5 months ago

Is contrastive divergence still useful for training or has it been supplanted by other methods?

permalink

[−] **andrewyng** [S] 13 points 5 months ago

In the early days of deep learning, Hinton had developed a few probabilistic deep learning algorithms such as Restricted Boltzmann Machines, which trained using contrastive divergence. But these models were really complicated, and computing the normalization constant (partition function) was intractable, leading to really complex MCMC and other algorithms for training them.

Over the next few years, we realized that these probabilistic formalisms didn't offer any advantage in most settings, but just added a lot of complexity. Thus, almost all of deep learning has since moved away from these probabilistic formalisms, to instead use neural networks with deterministic computations. One notable exception is that there're still a few groups (such as Ruslan Salakhutdinov's) doing very cool work on generative models using RBMs; but this is a minority. Most of deep learning is now done using backpropagation, and contrastive divergence is very rarely used.

As an aside, most of deep learning's successes today are due to supervised learning (trained with backprop). Looking a little further out, I'm still very excited about the potential of unsupervised learning, since we have a lot more unlabeled data than labeled data; it's just that we just don't know what are the right algorithms are for unsupervised, and lots more research is needed here!

permalink parent

    [−] **letitgo12345** 1 point 5 months ago

    Thanks! So are RBMs still the best for making generative models or even there auto-encoders, etc. are ahead?

    permalink parent

        [−] **alexmlamb** 2 points 5 months ago

        I think that variational autoencoders have been getting the best results for generative modeling.

        permalink parent

        load more comments (2 replies)

[−] **dylancaponi** 18 points 5 months ago

Hi Andrew, just finished your awesome Coursera class. Did you see my extra credit submission?

permalink

    [−] **eldeemon** 4 points 5 months ago

    lol this is great. reminds me of some of the videogrep demos.

    permalink parent

        [−] **dylancaponi** 3 points 5 months ago

        Thanks! That's what I used.

        permalink parent

[−] **icdwiwd** 6 points 5 months ago*

@andrewyng First of all I wanted to say that your ML course on Coursera was amazing. Thank you!

(1) How much learning others helped you to develop your own skills in ML? You definitely put a lot of effort to prepare your online materials. Do you do this only to help others or maybe while preparing your materials you have also learned a lot - for example maybe you often investigated some concepts more deeply than you knew them before only because you wanted to explain them to others as clearly as possible.

(2) You have both outstanding academic and commercial experience. Are there any ML concepts or intuitions which are easier or faster to learn when you work for companies? And inversely - are there things which are easier / faster to learn in the academic world? I'm asking because lot of ML engineers seems to have PhD. So how is it helpful? Are those paths (commercial vs academic) somehow different?

(3) Which set of skills you find the most important in the ML field - is it practical application of ML, statistics or maybe domain knowledge of a particular problem? For example lets assume that I want to develop a speech recognition system and I'm an expert in ML, but I do know nothing about audio processing. Do I have a chance to be successful?

permalink

[−] **andrewyng** [S] 8 points 5 months ago
Thank you for taking the Coursera ML MOOC!

(1) The old saw that teaching others helps you to learn really is true. FWIW though I think one of the reasons I've had a few successes in research is because I'm a decent teacher. This helps me to build a great team, and it's usually the team (not me) that comes up with many of the great ideas you see us publish and write about. I think innovation often requires the combination of dozens of ideas from multiple team members, so I spend a lot of time trying to build that great team that can have those ideas.

(2) A lot of deep learning progress is driven by computational scale, and by data. For example, I think the bleeding edge of deep learning is shifting to HPC (high performance computing aka supercomputers), which is what we're working on at Baidu. I've found it easier to build new HPC technologies and access huge amounts of data in a corporate context. I hope that governments will increase funding of basic research, so as to make these resources easier for universities all around the world to get.

(3) The skillset needed for different problems is different. But broadly, the two sources of "knowledge" a program can have about a problem are (i) what you hand-engineer, and (ii) what it learns by itself from data. In some fields (such as computer vision; and I predict increasingly so speech recognition and NLP in the future), the rapidly rising flood of data means that (ii) is now the dominant force, and thus the domain knowledge and the ability to hand-engineer little features is becoming less and less important. 5 years ago, it was really difficult to get involved in computer vision or speech recognition research, because there was a lot of domain knowledge you had to acquire. But thanks to the rise of deep learning and the rise of data, I think the learning curve is now easier/shallower, because what's driving progress is machine learning+data, and it's now less critical to know about and be able to hand-engineer as many corner cases for

these domains. I'm probably over-simplifying a bit, but now the winning approach is increasingly to code up a learning algorithm, using only a modest amount of domain knowledge, and then to give it a ton of data, and let the algorithm figure things out from the data.

permalink   parent

[−] **eftm** 3 points 5 months ago

When partitioning data into training/cross-validation/test sets, what do you think about the method of rotating which (one) example as the test set and using all the rest as the training set, computational constraints permitting?

And if you could cover one more class of algorithms in detail in your course, what would it be?

Thanks! Really enjoyed your coursera course.

permalink

[−] **solaris60sec** 4 points 5 months ago

What do you two consider to be the classic texts and papers of the ML field--the works that inspired you to begin and continue your research in ML? My background is in electrical engineering, and having just finished the Coursera ML course I would love to get more insight into how ML has evolved over time.

permalink

[−] **frankcarey** 5 points 5 months ago

In Andrew's talk at NVIDIA's GTC Conference, he makes a great analogy that DeepLearning is like building a rocket, where the engine is the network, and the fuel is the data, and either one can be a bottleneck.

What do you see as the biggest bottlenecks right now to making #DeepLearning even more robust beyond image and speech recognition? Is it finding employees, getting enough of the right data, GPU bandwidth, higher level tools, finding a the proper fit with new applications?

Andrew's talk: http://www.ustream.tv/recorded/60113824

permalink

[−] **jose_m_romero** 5 points 5 months ago

Thank you Dr. Ng for your ML Coursera class, this has helped me greatly on my career. Now for the questions:

1. What are the hottest research areas in ML & AI today? What are some of the short term technologies that need to be developed to help the discipline forward?

2. How different are the different Coursera Stanford courses from the ones offered to registered Stanford students?

3. What general guidance would you give someone who has been working in the tech industry for 20 years and wants to change careers paths into academia?

permalink

[−] **ralphplzgo** 2 points 5 months ago

Hello Andrew and Adam.

Thanks for coming and speaking at UC Berkeley awhile back. It was really interesting and inspiring. A few questions:

(1) As an undergraduate, after taking Cal's rough equivalent of CS229, what are the next steps in terms of becoming an ML expert in deep learning? Graduate courses? Getting involved in research? Self-studying certain texts?

(2) As someone who is soon to enter the workforce and is inspired by the technologies that have spawned out of Baidu's Inst. of Deep Learning and Google Brain, what are things that really makes someone stand out when hiring someone (new grad/intern) to work on these technologies? GPA? Publications? Side-projects? General coding proficiency?

(3) To Andrew - What were some of your favorite things to do or places to eat from your time at Cal?

Thanks for your time to do this AMA.

permalink

[−] **jmq1618** 2 points 5 months ago

Huge thank you to Dr. Ng, plus all TAs and others who made it possible, for the Coursera ML course. I don't know how I would have had access to such good guidance in this field without you! And thanks to both Dr. Ng & Dr. Coates for this AMA. :)

1. Having finished the course, I'm caught between the feeling of unlimited possibilities and having no idea where to start. Do you have suggestions for where newcomers can find datasets to practice and solidify the skills imparted by the Coursera course?

2. It seems linear regression addresses quantitative problems while logistic regression is for qualitative problems. Is this an accurate assessment? Either way, can you give a basic example of how one might address a problem with both elements? I'm thinking of say predicting a companies' revenues based on certain accounting metrics (quantitative) and market participation in certain product categories (qualitative). Please feel free to substitute a better example.

3. Let's be honest: what are the chances for someone breaking into this field who isn't in Silicon Valley? Looking at Prof. Ng's Stanford FAQ it's implied that only experienced individuals in the SF Bay area need apply. That's not a criticism but an observation. Am I wrong to assume that your paths to success are a) grind through competitions & academia until you get on with one of the ML "big boys" and/or a Silicon Valley start up or b) do your own thing (i.e. your own start up), solve a huge practical problem and hope you get exposure?

I've put my question in order of priority if there's insufficient time for all. Thanks again for your time & consideration.

permalink

[−] **deepc94** 2 points 5 months ago

Dear Prof. Ng, I have the following questions:

1) Can you tell us a little about convolutional neural netwoks and their application to computer vision? Also kindly suggest some resources (books/papers/web articles) on Deep Learning and Convolutional Neural Networks. :)

2) Can you suggest a comprehensive textbook to delve deeper into the more involved math

behind Machine Learning? Is PRML by C.M. Bishop a good choice? (and if so, what are your views about the pro-bayesian techniques in the book). Do you think non-parametric Bayesian methods for Computer Vision have a future?

3) Can you explain the necessity of tools such as Torch or Caffe?

P.S. I completed 100% of your ML course, and it was an amazing experience! Thanks a lot.

permalink

[−] **Foxtr0t** 2 points 5 months ago*

Professor,

1. What's your favourite *supervised* dimensionality reduction, or more generally - representation learning, method?
2. Why is there no mention of random forests in your ML MOOC on Coursera?
3. What's another word for pirate treasure?

permalink

[−] **yeehah** 2 points 5 months ago

Dr. Ng, how many students have taken your machine learning MOOC? I told my wife, who is also a professor, that you've had more than 100,000 students, but she's not convinced. How many students that sign up for the class actually complete the coursework?

permalink

[−] **massltime** 2 points 5 months ago

Hi, Prof Ng! I'm a undergraduate student in Beijing and studying in machine learning. Recently, I'm working on convolutional deep belief networks and applying it on signal processing. Could you please give me some advice on utilizing advanced feature learning algorithms(such ad CDBN) on signal processing? Finally, it's so excited to know that you have joined Baidu. I'm looking forward to do the research in Baidu with you after I graduate as Phd(maybe 6 years later~), or in summer internship :).

permalink

[−] **Zhiyang** 2 points 5 months ago

Hello Andrew! I have learned the Coursera Machine Learning Course since 2012. Recently we finished making the Chinese subtitles for this wonderful course. I was moved when listening to and translating your words in the last 7-min video of all the 113 videos. We just wonder if you could release some more advanced machine learning course or deep learning course so that we can have a much deeper understanding of this field, especially for us working in related fields? NOT in WIKI style, we want videos! We all love you and your teaching, and we are all looking forward to exploring your new courses! Thank you Andrew!

BTW, there are some mistakes in the 9-2 and 9-3 lecture videos, mainly about backpropagation. You missed some derivitive terms when deriving error delta terms in output layer and hidden layers. We compared it with your UFLDL Tutorial and found it might be your typo.

permalink

[−] **artificialcuriosity** 2 points 5 months ago

In your opinion, what are some promising research directions in building "deep" RL? (Besides recent approaches e.g. deep-Q network by DeepMind). Thanks!

permalink

[−] **Hwhacker** 2 points 5 months ago

Dear Professor Ng, Thanks for the ML course! It was interesting to get such an overview of techniques from Least Squares thought Neural Networks. The "unified treatment" of the subject brought out the interesting parallels between the different approaches. One comment - for myself I was able to produce an efficient vectorized implementation of each algorithm. But sometimes it felt like I only got the right answer through "dimensional analysis". (I.E., just make sure the dimensions in the Matrix calculations matched up). I felt at times that the code "wrote itself". And my comprehension of what I was doing lagged behind somewhat. My own picture of matrix multiplication is limited to "the repeated 'dot products' of the coefficients (aij) with the inputs (xi)". That was sufficient for most of the programming exercises. Except the last. The "collaborative filtering" exercise. I got the right answer here as well. But in the process of doing so I formed an "outer product" of all movies and features against the "thetas". And I really don't feel comfortable about my intuition as to what that "outer product" means. But its dimensions matched the Y matrix of current rankings, so I happily subtracted one from the other to get the ranking difference to work with. :-) But, again, I felt uncomfortable doing so, lacking a complete understanding. So, I find myself enrolled now in Dr Strang's Linear Algebra course to gain more insight. :-) But I wonder if you had any tips or suggested courses for curing my "matrix anxiety"? Jim

permalink

[−] **_caspinator_** 2 points 5 months ago

Hello Dr. Ng and Dr. Coates, do you have any thoughts on the feasibility of using ML to predict crimes? More specifically, given historic crime data for a region (complete with the GIS coordinate) would it be possible to predict how many crimes you would expect to see on a given day and where those crimes might occur? You could then use this data to plan patrol car deployments for instance.

Since the nature of the data is so rooted in human behaviour I'm wondering if that would cause problems for the ML algorithm. In your experience, do you have any insights in this area?

And thank you very much for the excellent Coursera course.

permalink

[−] **pranv** 6 points 5 months ago*

Hey,

- Pooling Operation for Downsampling seems to take away a lot of useful information while giving just a minor boost to spatial invariance. There is a paper that suggests using a convolution layer with large stride instead of pooling, and they even got better results with it. Have you guys thought of any method to eliminate or make pooling more efficient?

- You wrote a paper 4 years ago on using training methods other than SGD in Deep Networks like L-BFGS and Conjugate Gradient with Line Search, but yet, SGD is the most used method today. Do you think that there is a scope for these other methods? Why haven't they picked up usage even though you improved the state of the art

using these other methods?

- All machine learning methods go for correlation, and not causation. They are essentially designed to do that. Do you think that's a weakness in machine learning, that'll not go away? What can be done about it?

Thank You for taking time and doing this AMA. CHEERS!

permalink

[−] **Piximan** 4 points 5 months ago

This one is for Adam.

Your work that I'm most familiar with was exploring/describing single layer networks that performed better than the more complex/deep learning learning methods of the time on the CIFAR dataset.

Do you think that simpler configurations are possible that can compete with todays large network performance? Would it only be for certain dataset configurations that are difficult for large networks and their variants?

Thanks!

permalink

[−] **adamcoates** [Director of Baidu Research] 5 points 5 months ago

One of the reasons we looked at single layer networks was so that we could rapidly explore a lot of characteristics that we felt could influence how these models performed without a lot of the complexity that deep networks brought at the time (e.g., needing to train layer-by-layer). There is lots of evidence (empirical and theoretical) today, however, that deep networks can represent far more complex functions than shallow ones and, thus, to make use of the very large training datasets available, it is probably important to continue using large/deep networks for these problems.

Thankfully, while deep networks can be tricky to get working compared to some of the simplest models in 2011, today we have the benefit of much better tools and faster computers --- this lets us iterate quickly and explore in a way that we couldn't do in 2011. In some sense, building better systems for DL has enabled us to explore large, deep models at a pace similar to what we could do in 2011 only for very simple models. This is one of the reasons we invest a lot in systems research for deep learning here in the AI Lab: the faster we are able to run experiments, the more rapidly we can learn, and the easier it is to find models that are successful and understand all of the trade-offs.

Sometimes the "best" model ends up being a bit more complex than we want, but the good news is that the *process* of finding these models has been simplified a lot!

permalink   parent

[−] **elsonidoq** 1 point 5 months ago

Hi Adam! I have a follow up question regarding your answer. Do you have any recommended reading for the *process* of finding models?

permalink   parent

[−] [deleted] 5 months ago

[deleted]

[–] **NicolasGuacamole** 2 points 5 months ago

I'm confused as to what you mean by this, could you clarify it for me please?

permalink

[–] **xamdam** 5 points 5 months ago

Hi Andrew - huge fan of your ML course and Coursera in general - thanks!

My question is about the recent AGI safety controversy, particularly some quotes attributed to you here

http://fusion.net/story/54583/the-case-against-killer-robots-from-a-guy-actually-building-ai/

If I understand you correctly based on "overpopulation on Mars" statement you seem to be agreeing that human (and above) level AI can be dangerous (which seems logical) but disagreeing that it should be of current concern (I'm guessing overpopulation of Mars is hundreds of years away at least). Is that correct?

If so, what's your earliest estimate for such technologies to be developed? I realize there is a crapton of uncertainty, but I imagine you have some guesses.

Also, assuming safety research needs significant lead time to develop, when do you think it would be appropriate to start? How would we know? It seems like an important issue to get right.

Lastly, considering your own uncertainty, to what degree do you take other serious researcher's estimates into account? It seems people like Stuart Russell, Larry Wasserman, Shane Legg, Juergen Schmidhuber, Nils Nillson and Tim Gowers (granted not an AI researcher, but he worked a lot with theorem proving) have estimates of something like 50% chance of having human-level capability within 50 years.

Thanks a lot, looking forward to more great things from you!

permalink

[–] **AsIAm** 4 points 5 months ago*

Hey, Andrew and Adam!

- Andrew talked at GTC2015 about ML *"flight simulator"* that is helping novice Baidu engineers to experience many different training examples so they can learn the black magic tricks which are often used in NN community. Can you talk about the approach/tool more *concretly*?
- Do you see ML as the next programming paradigm that is going to dominate in the field of computer science?
- If you are going to do something like Google Brain but at Baidu, are you going to name it Chinese Room? ;)

Thank you for doing AMA!

permalink

[–] **LachineMearning** 2 points 5 months ago

If you would have 1000 times the memory (disk/ram) available compared to what you've used so far, what technique would become viable that is currently not, if any? What about 1000000?

If you would have 1000 times the processing power (in parallel) available compared to what

you've used so far, what technique would become viable that is currently not, if any? What about 1000000?

If you would have 1000 times the processing power (not in parallel, so pure speed/hz) available compared to what you've used so far, what technique would become viable that is currently not, if any? What about 1000000?

permalink

[−] **NicolasGuacamole** 1 point 5 months ago

ITT: Same techniques, bigger nets.

permalink   parent

[−] **NicolasGuacamole** 3 points 5 months ago

Do you believe that there is far more to be discovered in the area of regularisation? If so - why? And which regularisation techniques do you believe to be most promising?

permalink

[−] **lzpp** 4 points 5 months ago

Hi Andrew and Adam! Thanks for doing this!

1) How did you get into machine learning? That is, did you have an interest from a young age or did you get introduced in undergrad/grad school? In your opinion, is it a difficult or chancy path to take?

2) What in machine learning is now the "cutting edge"? Deep Learning? What specifically about it?

3) To Andrew specifically, how did you come up with Coursera and where do you see MOOCs going in the future?

4) How can I as a high school student advance my knowledge of machine learning? Machine learning is one of the more accessible fields of research (as opposed to, say, paleontology), but I still don't feel like I have a sense of where to go next. (Also, does Baidu offer internships? :))

permalink

[−] **zzerrp** 2 points 5 months ago

Hi Andrew, I have followed your work with interest and audited a few of your machine learning courses online. They have been an incredible resource. I actually made use of your homework exercises on the sparse autoencoder in my research on neural activity. So thanks for your dedication to education! I wanted to ask: when you are confronted with a large/high-dimensional/complex data set, what are the main early considerations that you use in determining what family of learning algorithms you will try with it? Do you have a recommended standard approach (e.g. start simple and linear and move to more complex techniques if those fail?) or are there things that you might notice in a data set that suggest that particular types of algorithms might be really well suited?

permalink

[−] **eftm** 1 point 5 months ago

He mentions in the Coursera course that he does start simple, plotting learning curves and such.

permalink   parent

[–] **llevar** 1 point 5 months ago

Pick the simplest model that can actually learn your function.

permalink parent

[–] **clbam8** 5 points 5 months ago

Thank you both for doing this AMA!

How do you guys compare your research work at Stanford University with your research at Baidu Lab?

permalink

[–] **NicolasGuacamole** 5 points 5 months ago

With some companies having such a great deal of capital (Baidu, Google et al.) experiments in industry can now be run using networks with a far greater number of parameters than one would see in academia.

Do you believe that for the foreseeable future Deep Learning / ML research will be done mostly in the private sector rather than in public institutions? What might the ramifications of this be?

permalink

[–] **sixwings** 2 points 5 months ago*

Thank you both for your time. Two questions for Dr. Ng. 1) Do you think AI research can continue to progress at a rapid rate without a transition from supervised to unsupervised learning? 2) Do you have a plan or a strategy on how best to achieve unsupervised learning?

permalink

[–] **tepsijash** 2 points 5 months ago

I would like to ask where do you see Machine Learning currently heading? I have the impression that it is growing in popularity, so I am interested about your thoughts on the current demand for data scientists?

Also for Dr. Ng, I have just finished your course on Coursera, and would like to thank you for this amazing resource. I am interested to know why is Reinforcement learning omitted from the course, and are you planning on creating more courses in the near future?

permalink

[–] **unital** 2 points 5 months ago

Hi Andrew and Adam, thanks for taking your time doing this. I also want to thank Andrew again for his great introductory ML course on Coursera.

My question is, how hard is it for a pure math major(or a pure math PhD) to break into the ML field(either industry or academia), if it's possible at all? Almost everyone in the field seems to have either a comp sci/ statistics or engineering background.

permalink

[–] **ford_beeblebrox** 2 points 5 months ago

What excites you about re-inforcement learning ?

permalink

[–] **gardinal** 0 points 5 months ago

Just finished my bachelors, how does 2 years of ML in the industry compare to 2 years of

Masters education in the relevant field?

permalink

[–] **jejonesphd**   1 point 5 months ago

What is the relationship of deep learning/sparse coding to what we have learned in the ML course?

permalink

[–] **mlhobbyist**   1 point 5 months ago

How might the major ML algorithms be complemented to perform better against time series feature data with seasonal variation?

permalink

[–] **lamecicle**   1 point 5 months ago

What can someone do to differentiate themselves in the field of competitive and research machine learning?

permalink

[–] **keoghpe**   1 point 5 months ago

As machine learning algorithms improve, do you see a place for knowledge based approaches in the future? Do you see any advantages to using argument based approaches, for example in expert systems?

permalink

[–] **Refefer**   1 point 5 months ago

What's your take on probabilistic programming languages in conjunction with how we conduct ML?

permalink

[–] **Power_Berries**   1 point 5 months ago

Andrew, what is the best way to spend the next 6-12 months advancing towards deep learning proficiency? I just completed your coursera course last night and have not yet implemented anything in the real world. Do I need to do this first?

permalink

[–] **overk4ll**   1 point 5 months ago

Hi, I'm wondering, what is your guys' take on neural tracing? Are you aware of projects focused on this outside of Seung Labs and Eyewire-collaborators? Would you have an idea of how soon we can expect the process to be completely automated?

permalink

[–] **onalurgidbee**   1 point 5 months ago

Do you think that GPUs will continue to be the best option for large-scale deep learning, or do you see it progressing to FPGAs (I know Microsoft has experimented with this) or even ASICs?

And thanks Andrew for the most recent Machine Learning course; I learned a lot and found it to be very well presented.

permalink

[–] **abreu0101**   1 point 5 months ago

Hi Andrew Ng/Adam Coates, thanks for the ML course. Some questions I have:

-What book of optimization, you recommend? -Mathematical topics which are important in the area of ML, you recommend? -How you prepare when studying a new topic? -Do You use a library ML developed by you? or use some library?

permalink

[−] **mllrnr** 1 point 5 months ago

First of all thank you for your course on ML. It was very informative. My question is how do you think students who are not based in the US should go about getting into this field after completing their bachelors, since there are still a lot of countries that don't have any good universities with ML courses.(sorry if this is too specific)

permalink

[−] **mszlazak** 1 point 5 months ago

@Andrew Ng. You recently mentioned in an interview that there was no killer app for machine learning went it comes to images (computer vision). I think killer app meant a market of at least 100 million people/users. Transportation is the obvious one that many would pick but if what's required is not all that sophisticated when it comes to image recognition then i see the problem. Care to elaborate more?

permalink

[−] **madvas** 1 point 5 months ago

Your and your team's work will probably liberate millions of people from their jobs in a near future. Unfortunately, current monetary social system requires people to go to their jobs and earn money constantly. Have you ever though about some social system, which might be a good alternative, when great amount of people won't have to work thanks to progress in AI and Machine learning?

permalink

[−] **jsflamholz** 1 point 5 months ago

How might one quantify a non-numerical feature like house-style {cape, colonial, ranch, mcMansion} in order to use, e.g., linear/logistic regression?

permalink

[−] **svanderbeck** 1 point 5 months ago

I love the advances in CNN's particularly for image classification tasks. However, having to define the network architecture still does not "feel" right in that defining the architecture is somewhat similar to defining features. Is it possible more generative networks are the next frontier in CNN's?

permalink

[−] **madvas** 1 point 5 months ago

In what year do you think we'll construct first neural net with number of connections/neurons equal to human brain?

permalink

[−] **drcleanser** 1 point 5 months ago

What is the status of ML being applied to biology/biotech?

I see tons of potential for using ML in drug design, crystallography, high throughput experiments, etc.... BUT a lot of examples of ML are focused on image, text or audio

analysis, primarily for social/civil projects.

I'm often irked by the amount of resources that could be going into doing incredible science and curing diseases, but instead end up in a toy/app/gadget or advertising...

permalink

[−] **pradyu1993** 1 point 5 months ago

Prof Andrew Ng What are your thoughts on Hierarchical temporal memory technology?

permalink

[−] **delude** 1 point 5 months ago

Long time admirer of your work and philosophies. Fascinated with machine learning and big data science in general. Studied programming and data mining in school and have taken your online machine learning class. Due to current life situation, am unable to move to west coast for career.

   1. What would you suggest as a next step for pursuing this passion?
   2. Favorite coding libraries and languages for machine learning?
   3. Most interesting personal discovery in your work?

permalink

[−] **jjm319** 1 point 5 months ago

you have mentioned a "flight simulator for research". The idea, as i understand, is to give students a lot of experience quickly. Would you expand on this idea? I imagine maybe testing research skills on known data sets. That is giving students data sets and asking them questions about it.

permalink

[−] **featalion** 1 point 5 months ago*

Drs. Ng and Coates, our brains are huge neural networks. And we are able to gather some information from it (EEG, MRI). Do you think, that it is possible "to decode" (say, classify) our thoughts, using such data and artificial neural networks? Thanks!

Many thanks to Dr. Ng for ML course on Coursera!

permalink

[−] **zweinstein** 1 point 5 months ago

Dear Dr. Ng and Dr. Coates, Would you please give suggestions on what are the next coursera classes (or other online resources) to take, after Dr. Ng's ML class, for beginners to become more proficient in ML? I understand it probably depends on the learning purpose. I want to use ML to handle some new metabolomics data, where the features are largely unknown. Thanks, Jen

permalink

    [−] **zweinstein** 1 point 5 months ago

    Dear Andrew, On a separate note, would you please tell us how to correctly pronounce your last name? :) Thanks! Jen

    permalink   parent

[−] **kp10g08** 1 point 5 months ago

Hi, thank you for the Coursera course on Machine Learning, I really enjoyed it and learn a lot! I have recently read on Probabilistic Programming and languages like Picture being

developed by MIT to tackle Machine Learning problems and make programming in this field more accessible. I was wondering what your thoughts are on the application of Probabilistic Programming and whether it will become a dominant part of Machine Learning? What types of applications can benefit from using PP?

permalink

[–] **TotesMessenger** 1 point 5 months ago*

This thread has been linked to from another place on reddit.

- [/r/artificial] AMA with Andrew Ng and Adam Coates
- [/r/remath] AMA Andrew Ng and Adam Coates : MachineLearning

If you follow any of the above links, respect the rules of reddit and don't vote. (Info / Contact)

permalink

[–] **filosoferthinker** 1 point 5 months ago

Will companies in Silicon Valley hire Developers for Machine Learning positions who are just Bachelors but have taken Machine Learning courses from Coursera or other MOOCs?

permalink

[–] **jkyle1234** 1 point 5 months ago

Hello Dr Ng do you see any lines of research for deep learning to incorporate a more online learning algorithm and getting rid of backpropegation as the method of training?

permalink

[–] **Ambiwlans** 1 point 5 months ago

How would you counter people saying that high accuracy image recognition cannot be done with an unembodied machine, or at least a machine that does not understand the environment that the photo represents. An example would be of identifying a black spec as a hockey puck because it is in a hockey game.

What is being done to solve this?

permalink

[–] **OdysseasB** 1 point 5 months ago

Dear Prof. Ng, Many thanks for the amazing Coursera Course on ML. Currently I am a PhD student on Computer Vision (Image Fusion) and would like to learn more advanced unsupervised ML algorithms (deep learning algorithms). Could you please recommend me some learning guidelines in order to advance my knowledge acquired from your Coursera course towards state of the art ML approaches applied in Computer Vision.

permalink

[–] **StargazyPi** 1 point 5 months ago

A question for you both: What problem excites you the most, that you think is "solvable" through Machine Learning within your lifetime?

Again, many thanks to you Andrew for the Coursera course (and for Coursera itself!) - you've inspired a great many people, myself included.

permalink

[–] **jjm319** 1 point 5 months ago

I have watched your image recognition research with interest. People mention the problems

with advanced AI but this i think is the most immediate one. Have you thought about how your research might be used by surveillance states?

permalink

[−] **artificialcuriosity** 1 point 5 months ago*

1. Can you elaborate more about pros and cons between undercomplete (bottleneck-based) and overcomplete (dictionary-based) representations, e.g. those used in autoencoders? Which one is better, in various possible senses, for example, efficiency, biologically plausibility, etc?

2. Most unsupervised learning (UL) algorithms try to minimize reconstruction error while incorporating built-in priors as regularizers, such as those discussed in Bengio et al's PAMI'2013 survey (smoothness, multiple explanatory factors, hierarchical organization of explanatory factors, shared factors across tasks, manifolds, natural clustering, temporal and spatial coherence, sparsity, simplicity of factor dependencies, etc.) Do you think "the one UL algorithm" that we are seeking may go beyond simply reconstructing inputs, say, to maximize some predictive information criteria? Would you like to elaborate more on some promising directions?

Thanks for sharing your precious time with us!

permalink

[−] **aiworld** 1 point 5 months ago

Given the central role that ImageNet has played in recent advances in ML, what would you like to see in future datasets and competitions to spur the next breakthroughs in ML?

permalink

[−] **Fledgeling** 1 point 5 months ago

Hello Dr. Ng,

I just recently finished your ML-008 course on coursera and I just wanted to say thank you for making it (and coursera) available.

I'm curious as to your thoughts on the future of machine learning in the Internet of Things space. It seems to me that the nature of the data sets themselves is shifting from a more centralized/focused set of data to a less focus/diversified set of sources and information types. We see social data, marketing data, location data, historical data, purchase records, weather data, etc. all getting lumped together and analyzed together.

Do you see any interesting trends in ML coming about due to this change in data trend?

permalink

[−] **bogdanmaksak** 1 point 5 months ago*

Andrew, firstly, huge thanks for your Coursera ML course and public talks you give!

My 3 questions are:

1. What do you think of graphical models/factor graphs, does it overlap with deep neural nets? Did you use factor graphs in any applications?

2. Do you use a particular library for deep learning prototyping/production, e.g. Torch, Theano, dl4j?

3. When and what was your first ML application?

permalink

[−] **mostlylearning** 1 point 5 months ago

In the introductory video, there was a teaser about using ML to separate 2 audio signals recorded on 2 microphones. I loved the course but am still itching to find out how to apply ML to this audio processing problem... please can you give me some pointers where to start? Thanks!! Ben

permalink

[−] **mszlazak** 1 point 5 months ago*

@Andew Ng. In this recent Re.Work 2015, you mentioned that you did not think the brain does back-propagation, that it works more like what people are trying to do in unsupervised learning. Why don't you think that brains do back-prop (and maybe delta-rule updating)? Also, experiments by Gallistel et. al. seem to suggest no delta-rule updating for certain things ("The Perception of Probability", 2014) but generally some version of that rule is quiet commonly used/assumed.Youtube here:

https://www.youtube.com/watch?feature=player_embedded&v=v5rEDe7Rwpg

permalink

[−] **bogdanmaksak** 1 point 5 months ago

Hi Adam

Which framework, e.g. Hadoop Map Reduce, Spark, did you use for distributing the tasks between CPUs/GPUs for training billions of connections size neural net?

permalink

   [−] **adamcoates** `Director of Baidu Research` 3 points 5 months ago

   Many of the current frameworks out there for large scale computation are very successful for problems involving huge amounts of data and relatively less computation. One of the things I worked on with Bryan Catanzaro and Andrew was how to do distributed computation for deep learning using tools/techniques that are specifically meant to handle very intense computational problems (like MPI/CUDA that come from the HPC/supercomputing world). There's more in our paper on that topic here: http://stanford.io/1JHzBwx

   Since a lot of the HPC tools ecosystem isn't as well developed for our problems, in the AI Lab / at Baidu Research the systems team is building a platform that let's DL researchers build experiments rapidly (like Hadoop/Spark do for cloud systems) but that run much faster!

   permalink   parent

   load more comments (1 reply)

[−] **shtankova** 1 point 5 months ago

1) What are the biggest mistakes that beginners make? What are the biggest wastes of time in learning process? 2) What do best ML and DM developers have in common? What skills and qualities do they possess? 3) Which of these skills and qualitites are not explained in courses and books? 4) What does the skill progression look like? What order is the best to develop profound knowledge of the field?

Thank you very much for your answers and for your course!

permalink

[–] **hgoldberg** 1 point 5 months ago

You've been able to widely disseminate foundations of ML as well as your particular practice through coursera. Have you noticed any influences or changes in the ML community that you might attribute to the availability of your course? Are you thinking of employing any ML techniques to use in teaching future versions of the course?

permalink

[–] **zevenozon** 1 point 5 months ago

I wonder why fuzzy logic is not covered in machine learning courses. It has a huge advantage over most other machine learning techniques in that rules obtained from 'experts' can easily be incorporated and used with those obtained using supervised learning, etc. We have used it successfully to solve problems in extractive metallurgy and business. I would like to hear your opinions on that.

permalink

[–] **Mahijoka** 1 point 5 months ago

Hi, Dr. andrew , I just finished Machine Learning course on Coursea.I have M.sc Computer science and strong mathematical background.I am implementing Naive Bay algo for categorisation of different web content in my selected categories list, e.g Gadgets, tech news, Data science etc.Will it be fine to implement unsupervised way, cluserting or supervised logistic regression way or combine naivies bay + logistic regression?

permalink

[–] **ras30** 1 point 5 months ago

What do you believe one should look for when hiring an ML researcher or ML engineer (since you mention that PhD is becoming lesser of a pre-requirement)? What do you look for at Baidu or have looked for in the past in your other positions?

permalink

[–] **gnarZeuce** 1 point 5 months ago

Hi Prof. Ng! As a big fan of your ML course on coursera, I was wondering what your suggestions for next steps would be to continually progress in the world of ML.

permalink

[–] **algol2000** 1 point 5 months ago*

Hi Prof. Ng., I am currently enrolled in your machine learning course and I am enjoying it! Thank you for making it available. There is a problem domain involving the identification and tracking of cars, in order to evaluate the speed of traffic. I was wondering, have you ever been involved with this kind of problem space?

permalink

[–] **heartzealhere** 1 point 5 months ago

Dr Ng,

thx for the ML course, I thoroughly enjoyed this course. My question - Are there any Open source or other ML initiatives/projects?

Are there any Open source projects or any other ML Projects that Dr. Ng or his team are driving? It would be a good validation of our learnings to partake and contribute to such

projects.

I would love to join such an effort part time.

thx

permalink

[−] **jollybobbyroger** 1 point 5 months ago

I am looking forward to taking the Coursera class on machine learning, but I was always disdained by the simple fact that the chosen language for the course is Matlab/Octave. I know how highly regarded this class is and will ignore my qualms with Matlab/Octave, but I would really like to know the reason for why this language was chosen over Python, which is a really pleasant language to use and learn.

Thank you for doing the AMA.

permalink

[−] **jmq1618** 1 point 5 months ago

I recall several references to Python, C++, Java, in the Coursera course. I imagine a lot of the big data sets out in the world to be SQL driven, and a quick Google shows some in pursuit of SQL based algorithm solutions.

Can you please comment on some practical considerations regarding algorithm implementation to scale? I'm thinking of a comment about Python linear algebra libraries that are built to utilize multiple-cores/machines. Not asking for a "best language" per se. I think a process for dropping formatted data files for Octave to run is fine for us just learning to learn, but switching to production languages seems useful to keep in mind.

Moreover, do you see any difference in opportunities between analyzing preexisting data (old big data) versus developing information gathering systems with ML algorithms in mind? Concretely, is the data gathering role of a ML developer greater than, less than or equal to the role of utilizing data in learning algorithms? It seems to me conversions of data others gather into useful matrices, rather than directing the gathering itself, predominates. I'd be very interested in your perspective as no one will hire me for machine learning but I have a lot of experience in data analysis and am hoping one skill set will dovetail into the other. :)

permalink

[−] **Razorwindsg** 1 point 5 months ago

Hi Andrew,

I see that there are still many companies who are stuck in Excel dashboards (means, compare week to week etc). While business reporting is still essential, how can we move employees from an "Excel" mindset to an "database" mindset that is required in predictive analytics?

Is it likely that we will see an affordable and easy to use machine learning package that the usual office worker can use? While we don't expect the receptionist to start predicting traffic flow, how far are we from the day that "normal" employees squeeze predictive analytics to it's limits? (much like how Excel is worked to its limits in most of today's business intelligence context)

There is a huge amount of papers written for algorithms and network architecture. Most

businesses seem to however fumble and still get lots of dirty data that is as of result of improper data design, or that they have data structures that isn't optimized for the pulls they need. Will we see a unified theory for how common businesses should structure their databases for common analytic tasks?

permalink

[–] **lczarne** 1 point 5 months ago

What do you think about doing Coursera's Johns Hopkins Data Science Specialization (9 courses) If someone is interested mainly in Machine Learning?

permalink

[–] **jeanlov** 1 point 5 months ago

I am really like this area of machine learning, but in this moment i am studying a part of information criteria like AIC(Akaike information criteria) that involve part modeling, statistics and information theory in systems identification for my master research. Is that you know of some interaction of this area and machine learning?. i really would like to know that. Thank you for you answer Dr. Ng.

permalink

[–] **freexstate** 1 point 5 months ago

Thank you Dr. Ng for your Coursera Machine Learning class. I teach Calculus and Statistics myself. What advice do you have for converting to a hybrid class format, making videos, and/or creating a Coursera course?

permalink

[–] **rakesh_pandey** 1 point 5 months ago

Prof. Ng you have mentioned at a place in this AMA: " the mission of the Baidu's AI Lab is to develop hard AI technologies that let us impact hundreds of millions of users." Can you please give a few examples of such potential technologies and their application. Thanks.

permalink

[–] **solsav** 1 point 5 months ago

Hi Andrew, thanks for taking the time. I would like to know if you reckon ML skills that you seek when hiring people in Baidu as talent or something an ordinary college students can acquire via hard work.

permalink

[–] **phoenixkbb** 1 point 5 months ago

[THANKS for ML on Coursera and Question about Baidu IDL]

- Hi Dear Prof. Andrew Ng, can't thank you and your team any more for the amazing ML course on Coursera. It made me more enthusiastic with data mining and machine learning and more determined with the career direction. :P

- I have some questions for Prof. Ng and Dr. Coates about Baidu deep learning institute. From the website of IDL Baidu, it seems that the main research area IDL now focusing on are computer vision and deep learning. Is there any projects in IDL related with text mining and user behavior analysis? Would there be any job opportunities for researchers now dealing with recommender system and topic model in IDL? Is there any hard requirement if we want to apply a job in IDL? Thank you so

much. :P

- It would be sincerely appreciated in case you could answer the questions. Have a nice day and kind regards!:)

Jing Yuan, 14.04.2015

permalink

[–] **raulpopa** 1 point 5 months ago*

Prof. Ng, you're THE best!

First question: Can you tell me a desired average error in linear and logistic regression? What percentage of error and what amount of false positive/negative results are ok in such algorithms? When the training set is quite low (500 items) how do you solve over-fitting? How do you validate the algorithm? Do you split your training set in half (one for learning, the other for validating), or you keep about 100 or less items for validation? Can you recommend a best practice?

I've also played a bit with GPA (Google Prediction API) and sent the same data I've used for my linear regressions (solved with normal equations on my side). I've linearized some features before (I've used a linearized versions f(x) of some features x) and then plugged the data through both my Octave code for prediction with normal equation and separately through GPA. The results were almost identical with minor differences. The average error I've got for my main project (about 100 features, and thousands of items), was 10.3 with normal equation and 10.4 with GPA. Plotting the data revealed that GPA solved the problem a bit differently.

Second question: Do you happen to know and can explain how does GPA work and why it tries a different route (possible some gradient descent or so) when all features are linear and when normal equation can solve this type of problem best anyway. Is it that GPA tries to reduce some over-fitting, which results in differences and higher error?

permalink

[–] **seeker_factory** 1 point 5 months ago

Question: Difference between Learning(Supervised.Classification and Unsupervised.Associative)

Both the techniques require learning that has data coming from previous experience(hence forming Association) and creates any room to identify the new requirement in these cases. Normally Unsupervised Learning doesnt need labeled data but when Association is created it automatically is labeled(hence confusion).

(sorry was lost identifying correct thread :(and had no clue how much time you guys will be on them, hence deleted twice before posting here)

permalink

[–] **halfourname** 1 point 5 months ago

I have found lots of material about designing, training and optimizing deep learning but there seems to be a real void around what happens with fully trained sets of weights and methods.

Do you know if anyone is archiving these or if there is a move on to reuse or extract information from already trained sets of weights and the networks that use them?

It seems like an under-studied area. Thanks for the AMA!

permalink

[–] **databoydg** 1 point 5 months ago

Depth vs. Breadth As someone with intermediate experience in ML (Master's AI/CV), I often find myself at a crossroads of how to best proceed. I am often daunted by the size of the field, and how much of the field I don't fully understand. In terms of furthering my expertise in the field, what are your thoughts on the tradeoffs of going deeper within one specific subfield, vs establishing a stronger foundation within the general field?

permalink

[–] **hvn92** 1 point 5 months ago

Hi Professor Ng,

I am a Computer Science student starting to learn about ML. I am taking your Coursera course and I am really enjoying it! What ML books would you suggest to beginning learners like me?

permalink

[–] **Fortyq** 1 point 5 months ago

Hi, Ng! When strong AI will arrive, do you think current deep learning models will have something to do with that? Will it be some crazy scaled version of deep NN, or some new kind of model?

permalink

[–] **saikrishnar** 1 point 5 months ago

Sir,

What are the features you're using in Deep Speech??

permalink

[–] **aditya_arun** 1 point 5 months ago

Hi Andrew and Adam, thanks for taking out time for this.

1) How do you see the area of unsupervised, semi-supervised, weakly supervised learning evolving and shaping the area of deep learning respectively given that lot of work these days are about getting lot of labelled data, and building larger models with nice initialization of parameters. How is industry moving forward with these areas?

2) Do you see a unified pattern recognizing algorithm that can perform many "complex higher order task" with little or no modification coming out any time soon? Do you think such a model is possible/feasible?

3) What would be your suggestions to a person (like me :P) who has just started his work in the area of ML/Deep Learning?

permalink

[–] **GnomyGnomy7** 1 point 5 months ago

Wow, Andrew NG, thanks a lot for taking your time here on Reddit, and for introducing me to ML. I am indeed indebted to you. I enjoyed your course on Coursera thoroughly, and literally shed a small tear, on the final week of the course.

My question is, throughout the course, you were encouraging us by saying that we are

catching up with the professionals in the field. Each time I heard you say that, I felt a lot more confident, and developed more interest towards the subject.

So how true were your words, and how much of it was to encourage us?

permalink

[–] **qnoxo** 1 point 5 months ago

Hello Andrew Ng! It is a big plesure to have you around here. My question is related to the results of a recent practical paper comparing loads of classifiers (174) against 120 different UCI datasets [1]. Based on this experiment, they conclude that the classifiers most likely to be the bests are the random forest (RF) versions, the best of which (implemented in R and accessed via caret) achieves 94.1% of the maximum accuracy overcoming 90% in the 84.3% of the data sets. Have you got the opportunity to take a look to this work? in that case, what do you think about the results? are those consistent with your experience?

Thank you!

[1] http://jmlr.csail.mit.edu/papers/volume15/delgado14a/delgado14a.pdf

permalink

[–] **awhitesong** 1 point 5 months ago*

Hi Adam and Andrew,

A lot of work has been done in neural networks and they are being avidly applied in a lot of areas. But I've seen somewhat less development in Spiking Neural Networks considering their great potential in robotics and relatively other fields. What do you think is the potential of Spiking neural networks and the scope of its development in terms of applications and research in the coming years?

PS: I am your coursera class's ex-student working on a research in SpikingNN seeing its great potential, and wanted an industrial viewpoint on this. So wanted to know this for a long time, just could't get to the AMA at the right moment. :)

permalink

[–] **nlpkid** 1 point 5 months ago

Which approach to AI do you think will be more powerful? The neuroscience and biology heavy method, or the more statistics oriented approach?

permalink

load more comments (67 replies)

---

about

blog
about
values
team
source code
advertise
jobs

help

site rules
FAQ
wiki
reddiquette
transparency
contact us

apps & tools

Alien Blue iOS app
reddit AMA app
mobile beta
buttons

<3

reddit gold
reddit store
redditgifts
reddit.tv

π