Presidential Address: Machine Learning For Finance

Murray Zed Frank
(Minnesota, and SAIF)

2018 Midwest Finance Association Meetings

San Antonio, Texas

“**AI is probably the most important thing humanity has ever worked on. I think of it as something more profound than electricity or fire**.”
Google CEO Sundar Pichai (January 2018)

Even if his claim is exaggerated, the impact over the next decade or two seems likely to be at least as big as the PC or the Internet – and you never know, the CEO of Google just might be right.

Significant change is already underway. It will affect our research, our teaching, and the demand for our services.

Plan for today

1. What is machine learning (ML)?
2. The ML explosion underway
3. Some important ML uses
4. ML impact on finance
5. Conclusion/some resources

Part 1.
How Different is Machine Learning Really?

1. Statistics as reinvented by applied computer scientists
   - Very heavy focus on correct out-of-sample predictions (to avoid overfitting)
   - Limited attention to standard errors and causality
2. Can also be viewed as a new approach to writing programs. Karpathy (Director of AI at Tesla) calls this Software 2.0. https://medium.com/@karpathy/software-2-0-a64152b37c35

- Applied perspective: If the algorithm actually works, then all is well.

- We need more on causality, standard errors, “why”, “understanding”
  - To justify medical treatments, protection against lawsuits, answer regulators, etc.

Machine Learning Types of Algorithms

- Supervised learning
  - start with examples that have inputs correctly matched to particular outputs. Then generalize to other examples that have inputs, but no output measures provided.
  - Where recent major advances have happened: deep learning, boosting, lasso
- Unsupervised learning
  - data reduction and summarization, finding natural groups or patterns, anomaly detection
  - Less developed so far: principal components, hierarchical cluster
  - In a sense, this is what much human learning is like – experiential
- Reinforcement learning
  - Approximate dynamic programming – close to economics and finance
  - Used in game playing (e.g. Alpha Zero)
- Generative Adversarial Nets (Goodfellow et al 2014)
  - Two networks (agents) interact to produce an outcome; often a zero-sum game
  - Agent 1 tries to max agent 2’s error rate, and agent 2 tries to min the error rate
  - E.g. make a photo that cannot be distinguished from natural photos

Typical ML Steps

1. Divide the data into training, validation and testing sets (75,15,10)
2. Regularize: a modification intended to reduce the testing error but not the training error
    - Dropout, penalty for complexity, etc.
3. Use the training set to estimate parameters
    - Assumes given hyperparameters
4. Pick hyperparameters that produces parameters that work on the validation set, or pick by cross-validation
5. Then evaluate the final model on the testing set

Why Do Deep Networks Work?

- We do not really know why

- Deep networks commonly have far more model parameters than the number of samples used for training. Yet some of these models have surprisingly small difference between "training error" and "test error"
  - Why doesn't the algorithm just memorize the data in sample?

- Leading attempts to explain why, all seem to fail when tested carefully (Zhang, Bengio, Hardt, Recht and Vinyals, 2017)
  - "when trained on a completely random labeling of the true data, neural networks achieve 0 training error. The test error, of course, is no better than random chance as there is no correlation between the training labels and the test labels."

The Unreasonable Effectiveness of Data

- As the data grows very large almost any "reasonable" algorithm will work fairly well
  - Halevy, A., Norvig, P., and Pereira, F. (2009) The unreasonable effectiveness of data. *IEEE Intelligent Systems*, *24*,2, 8-12.
- Example: de Fortuny, Martens, and Provost (2013) Predictive Modeling with Big Data: Is Bigger Really Better? Big Data, 1,4, 215-26.

TABLE 1. SPARSE DATASETS CONSIDERED

Dataset	*Active elements*	*Instances*	*Features*	*Sparseness*	*Target variable*
Yahoo Movies	220,831	7,620	11,914	0.99756753	Gender
Book Crossing	680,194	61,308	282,700	0.999960755	Age
Ta-Feng	723,449	31,640	23,718	0.999035964	Age
Dating	17,359,100	135,358	168,790	0.999240205	Gender
URL	277,058,644	2,396,130	3,230,441	0.999964207	Malicious or not
KDD-A	305,613,510	8,407,752	19,306,082	0.999998117	Correctness of answer
KDD-B	566,345,888	19,264,097	28,875,156	0.999998982	Correctness of answer
Flickr	34,645,468	11,195,143	497,470	0.999993779	Comments (few or many)
Bank	20,914,533	1,204,727	3,139,575	0.99999447	Product interest

The table reports the total number of active (non-zero) elements, the number of data instances, the number of features, the resultant sparseness for a traditional data matrix representation (fraction of matrix elements that are not active), and the target variable.

de Fortuny, et al (2013)

- Naive Bayes classification on several online content providers and a bank
- Performance continues to improve (some diminishing returns), even with number of observations gets huge

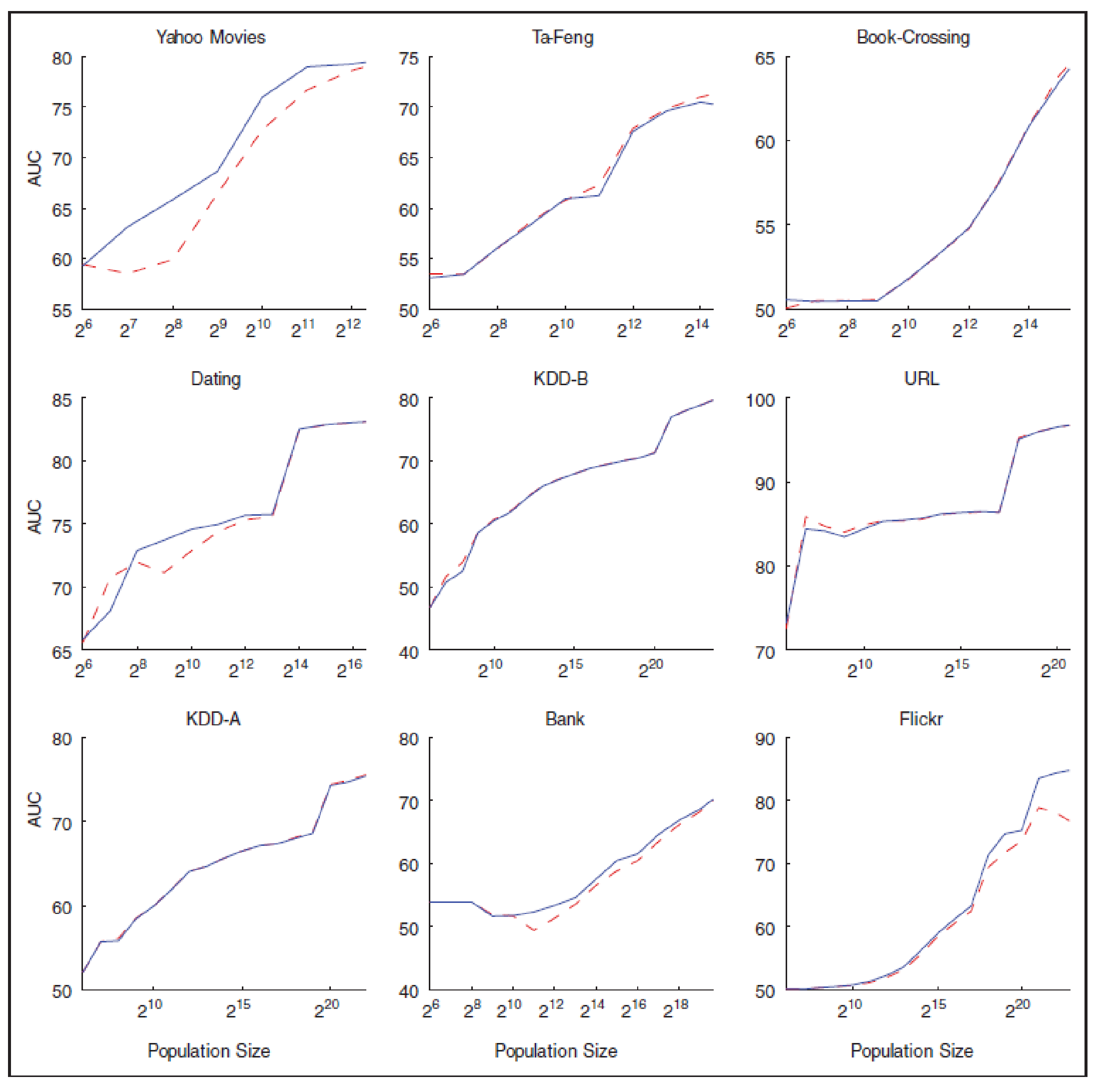


FIG. 2. Learning curves for the multivariate (solid, blue) and multinomial (dashed, red) Naive Bayes variants. For all of the datasets, adding more data (x-axis) leads to continued improvement of predictive power (area under the curve [AUC]).

No Free Lunch Theorem

- Wolpert (1996) “The Lack of A Priori Distinctions Between Learning Algorithms” Neural Computation.
  - If you are unwilling to make any assumptions then you have no reason to prefer one algorithm to another, until you try.
  - Some models are linear, others are not. Some might fit neural nets well, others might fit SVM or Naive Bayes. No model is a priori guaranteed to work better.

- We are forced to be “reasonable” in picking models to consider.

Part 2.
Happening Now on a Large Scale

- Vast investments underway (many $ billions annually) in chip improvements, in improving algorithms, and in adopting ML methods.
  - More advances will inevitably take place in the next few years

- Accenture Research spoke with more than 1,200 CEOs and top executives about AI in January 2018
  - "Three quarters (74 percent) of executives say they plan to use AI to automate tasks to a large or a very large extent in the next three years. But almost all (97 percent) note they intend to use AI to enhance worker capabilities"
  - 75% are currently accelerating investments in AI
  - 72% say this is a response to a competitive necessity

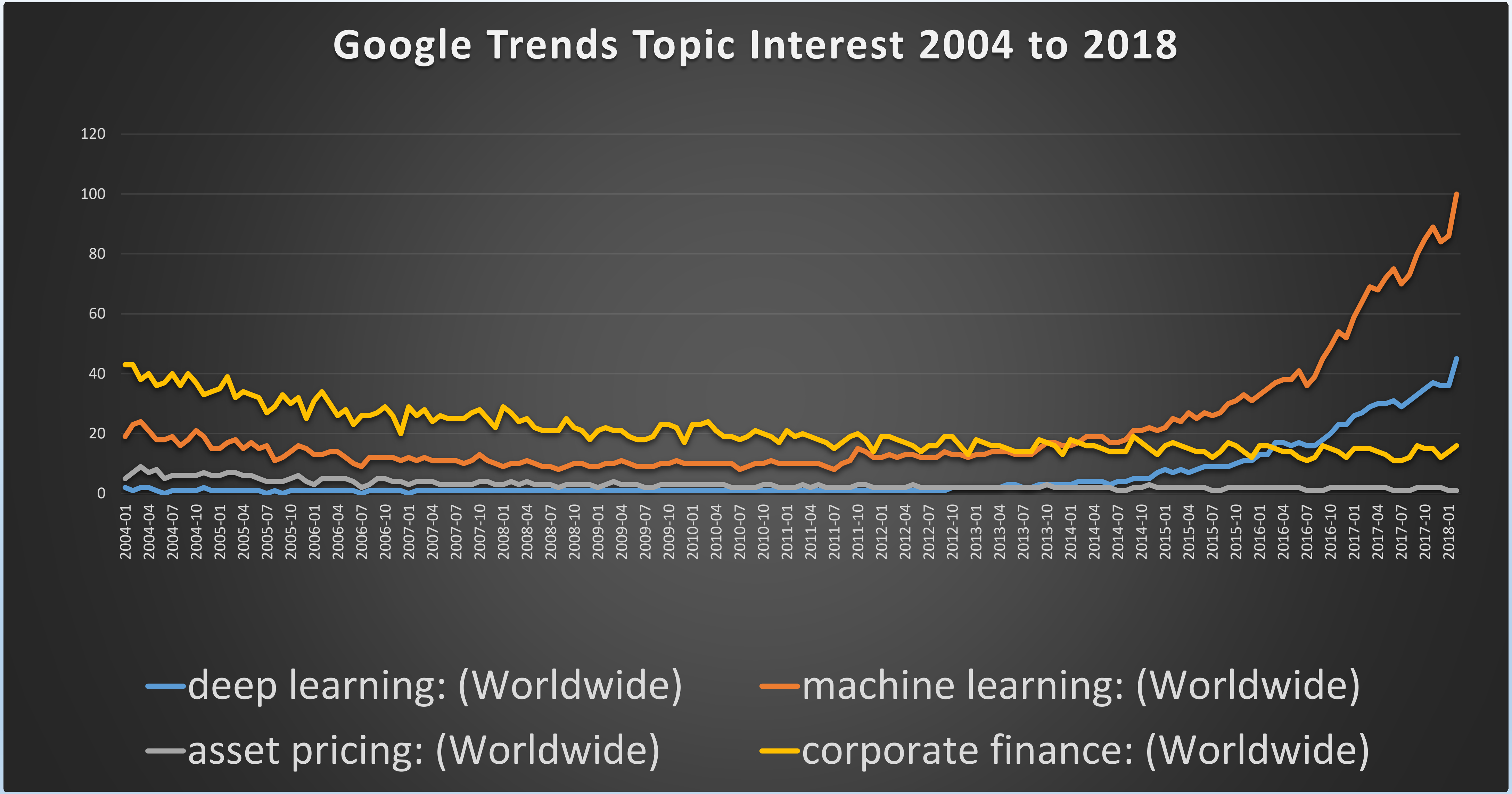
- "Microsoft has invested heavily in artificial intelligence, with its AI and Research group formed by Microsoft CEO Satya Nadella in 2016 as a fourth engineering division at the company. In its first year, the group grew by 60 percent to more than 8,000 people." https://www.geekwire.com/2018/microsoft-alibaba-ai-programs-beat-humans-stanford-reading-test-1st-time/
  - **How many academic finance departments would we need to get to 8,000 people**?

Google Scholar, Number of Search Results

(Feb 7, 2018)

- Finance: About 3,290,000 results
  - Asset pricing: About 1,720,000 results
  - Banking: About 2,920,000 results
  - Corporate Finance: About 2,310,000 results
  - Option Pricing: About 2,040,000 results

- Classification Trees: About 2,880,000 results
- Deep Learning: About 4,100,000 results
- Machine Learning: About 3,970,000 results
- Random Forest: About 2,320,000 results
- Reinforcement Learning: About 2,300,000 results
- Support Vector Machine: About 2,580,000 results


Google Trends Topic Interest 2004 to 2018
120
100
80
60
40
20
0
2004-01
2004-04
2004-07
2004-10
2005-01
2005-04
2005-07
2005-10
2006-01
2006-04
2006-07
2006-10
2007-01
2007-04
2007-07
2007-10
2008-01
2008-04
2008-07
2008-10
2009-01
2009-04
2009-07
2009-10
2010-01
2010-04
2010-07
2010-10
2011-01
2011-04
2011-07
2011-10
2012-01
2012-04
2012-07
2012-10
2013-01
2013-04
2013-07
2013-10
2014-01
2014-04
2014-07
2014-10
2015-01
2015-04
2015-07
2015-10
2016-01
2016-04
2016-07
2016-10
2017-01
2017-04
2017-07
2017-10
2018-01
deep learning: (Worldwide)
machine learning: (Worldwide)
asset pricing: (Worldwide)
corporate finance: (Worldwide)

In 2012 the World Changed

- Something of fundamental importance happened in late 2012 that is in the process of changing the world
  - What happened?
  - Why did it happen?
  - How is this affecting finance practice already?
  - How is it likely to affect academic finance over the next decade?

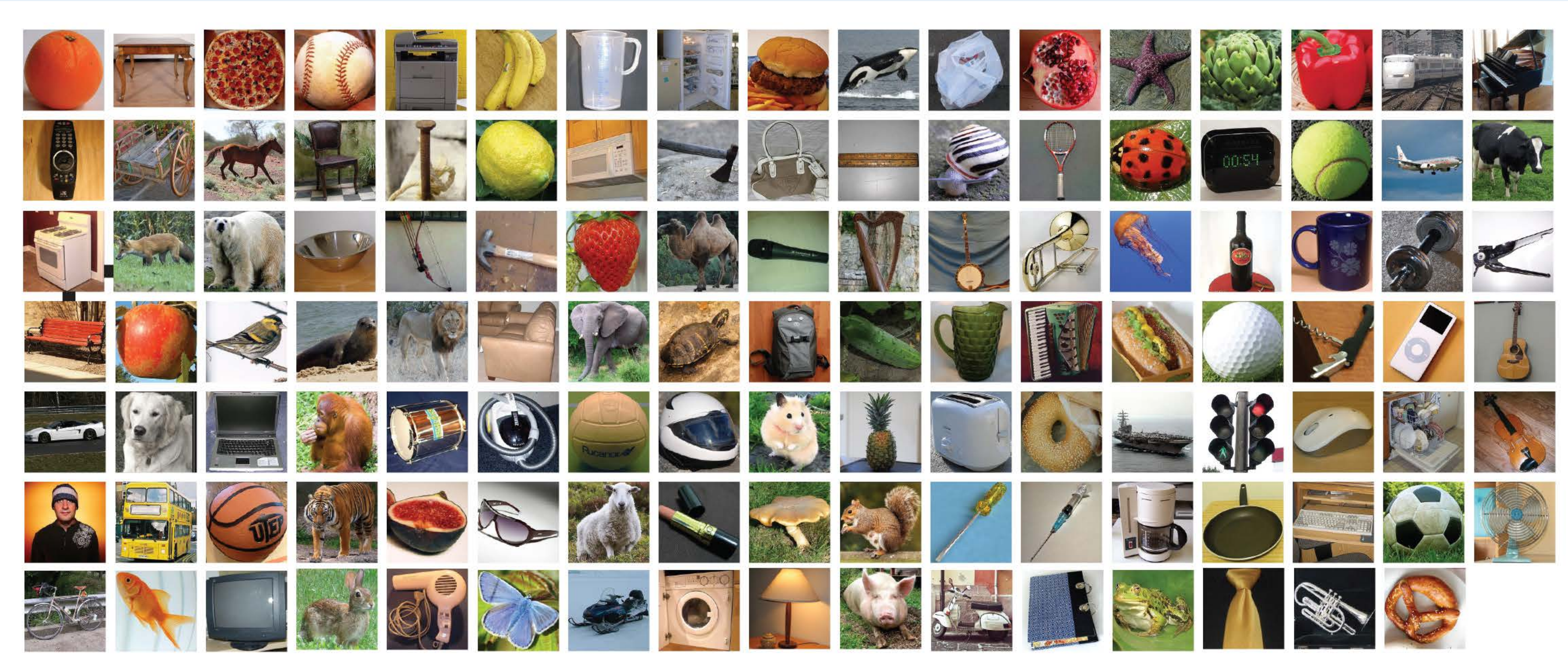
- Change is happening, and the impact may be large
  - Impact on scholarly methods
  - Impact on teaching methods
  - Impact on the demand for our students

What Happened?
The 2012 Imagenet Contest

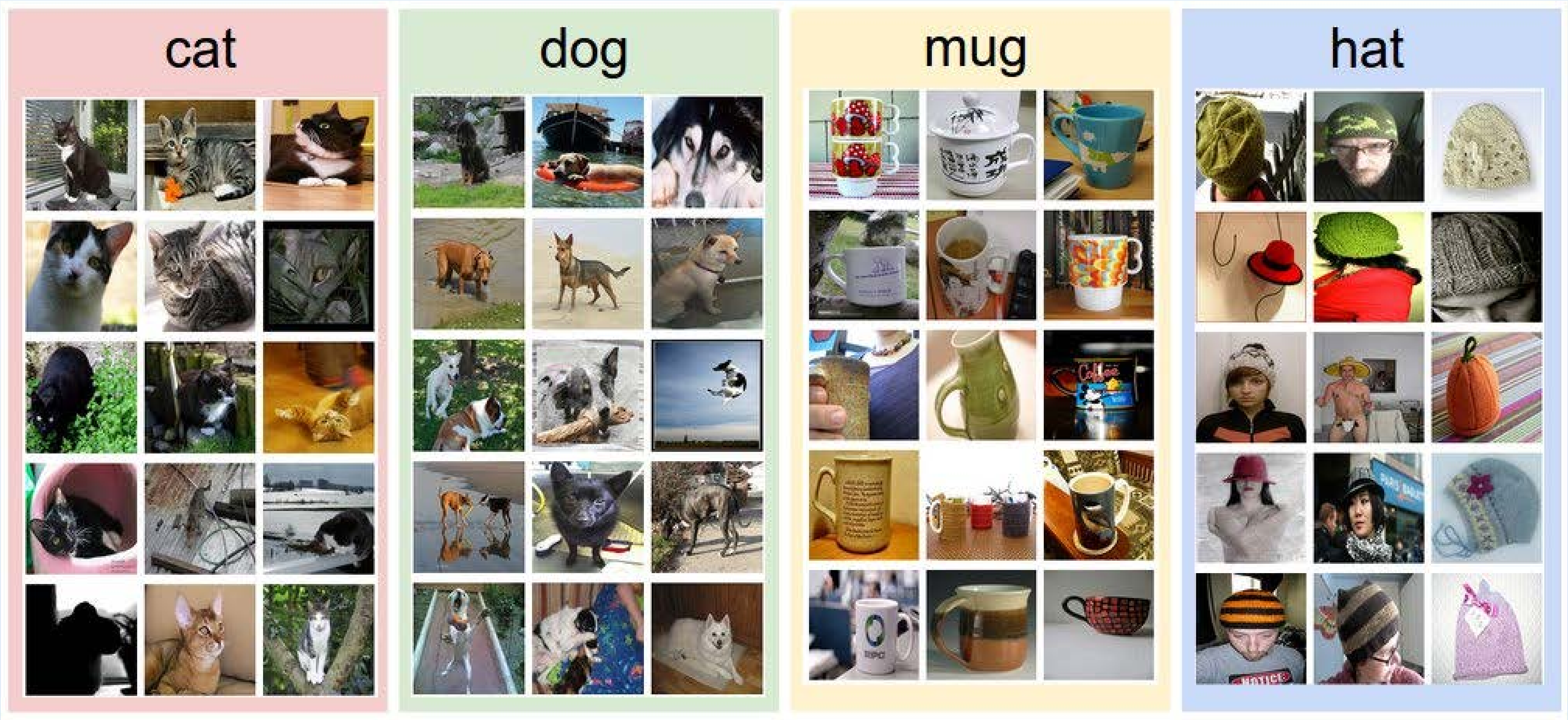
- Contest has 1.2 million training images and 50,000 validation images.
- Goal: to classify 150,000 images into 1000 categories.

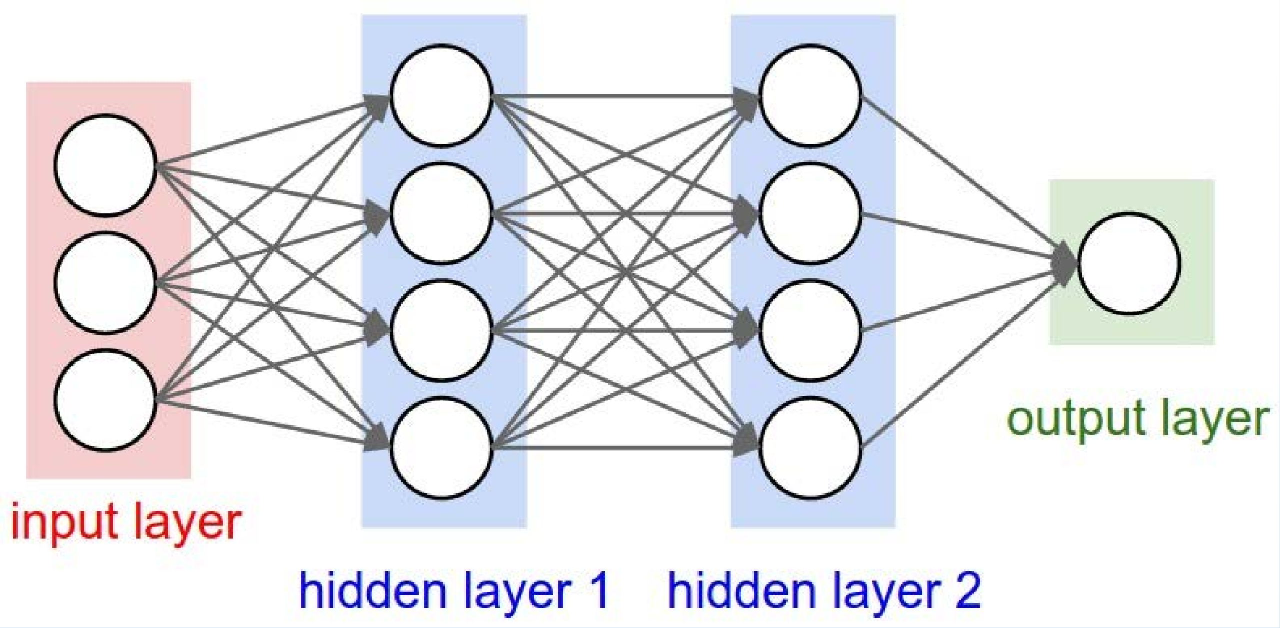
- Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems,* pages 1097-1105.
  - Large training dataset, NVIDIA GPU, convolutional deep neural net, and then … it works!
  - 60 million parameters to estimate in 5 convolutional layers and 3 fully connected layers
  - This 5 year old paper has 19,531 cites on Google Scholar (Feb 7, 2018)

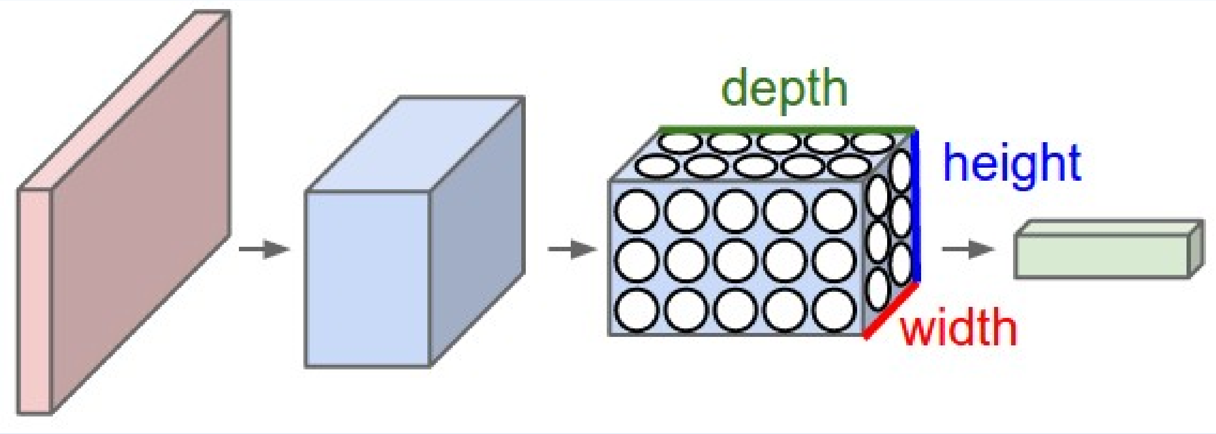
- Some famous finance papers
  - Modigliani Miller (1958 AER) has 21,481 cites
  - CAPM: Sharpe (1964 JF) has 20,767 cites
  - Efficient Markets Hypothesis: Fama (1970 JF) has 20,918 cites
  - Option pricing: Black Scholes (1973 JPE) has 35,186 cites


00:54

Examples of Training Labels







Left: A regular 3-layer Neural Network. Right: A ConvNet arranges its neurons in three dimensions (width, height, depth), as visualized in one of the layers. Every layer of a ConvNet transforms the 3D input volume to a 3D output volume of neuron activations. In this example, the red input layer holds the image, so its width and height would be the dimensions of the image, and the depth would be 3 (Red, Green, Blue channels)

http://cs231n.github.io/convolutional-networks/

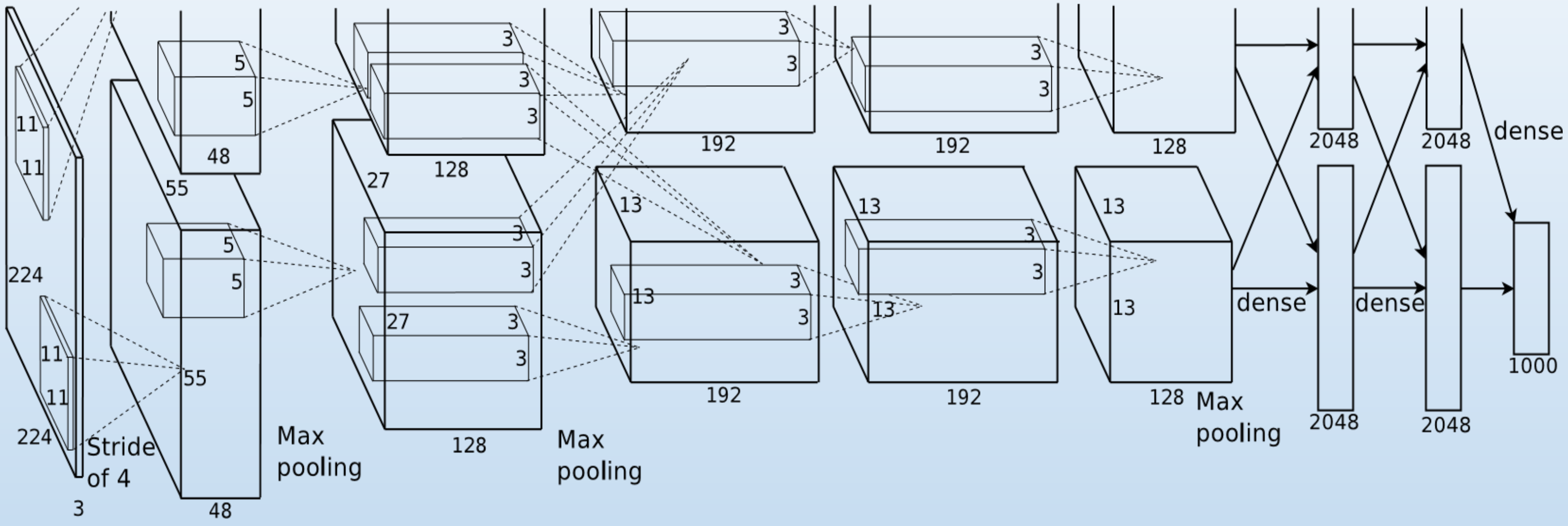


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input is 150,528-dimensional, and the number of neurons in the network's remaining layers is given by 253,440–186,624–64,896–64,896–43,264– 4096–4096–1000 (Krizhevsky et al 2012)

2012 Results

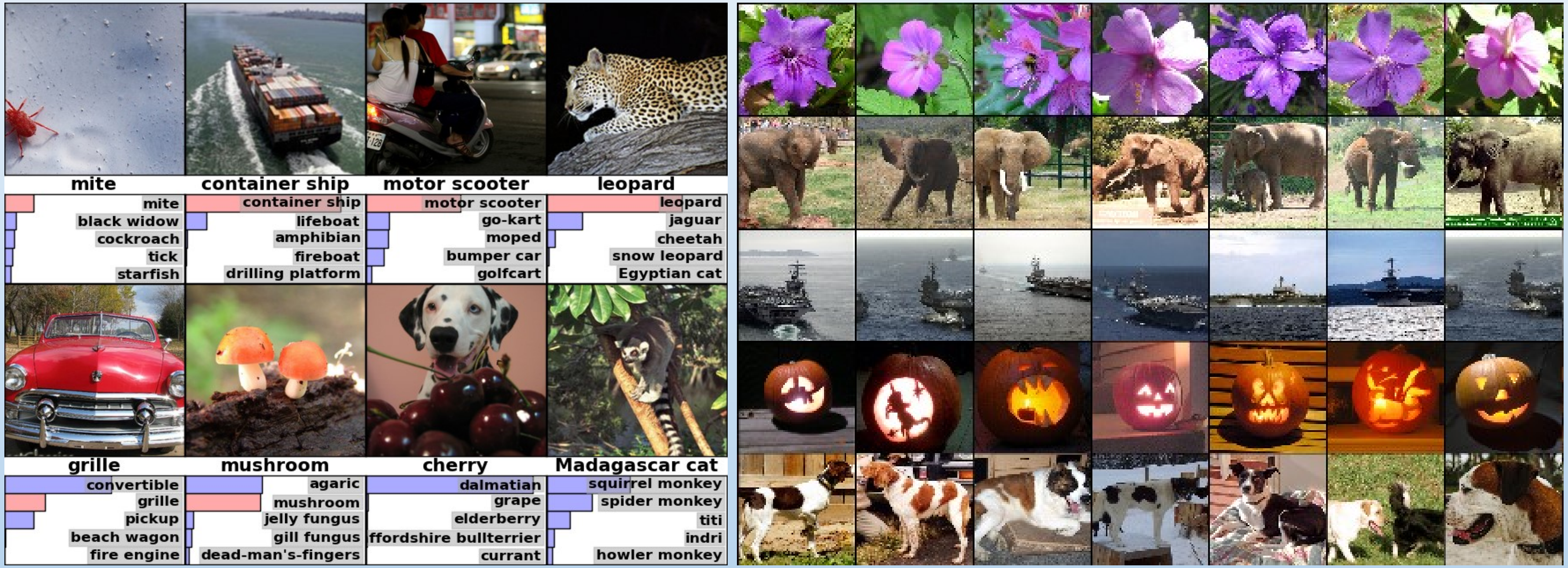
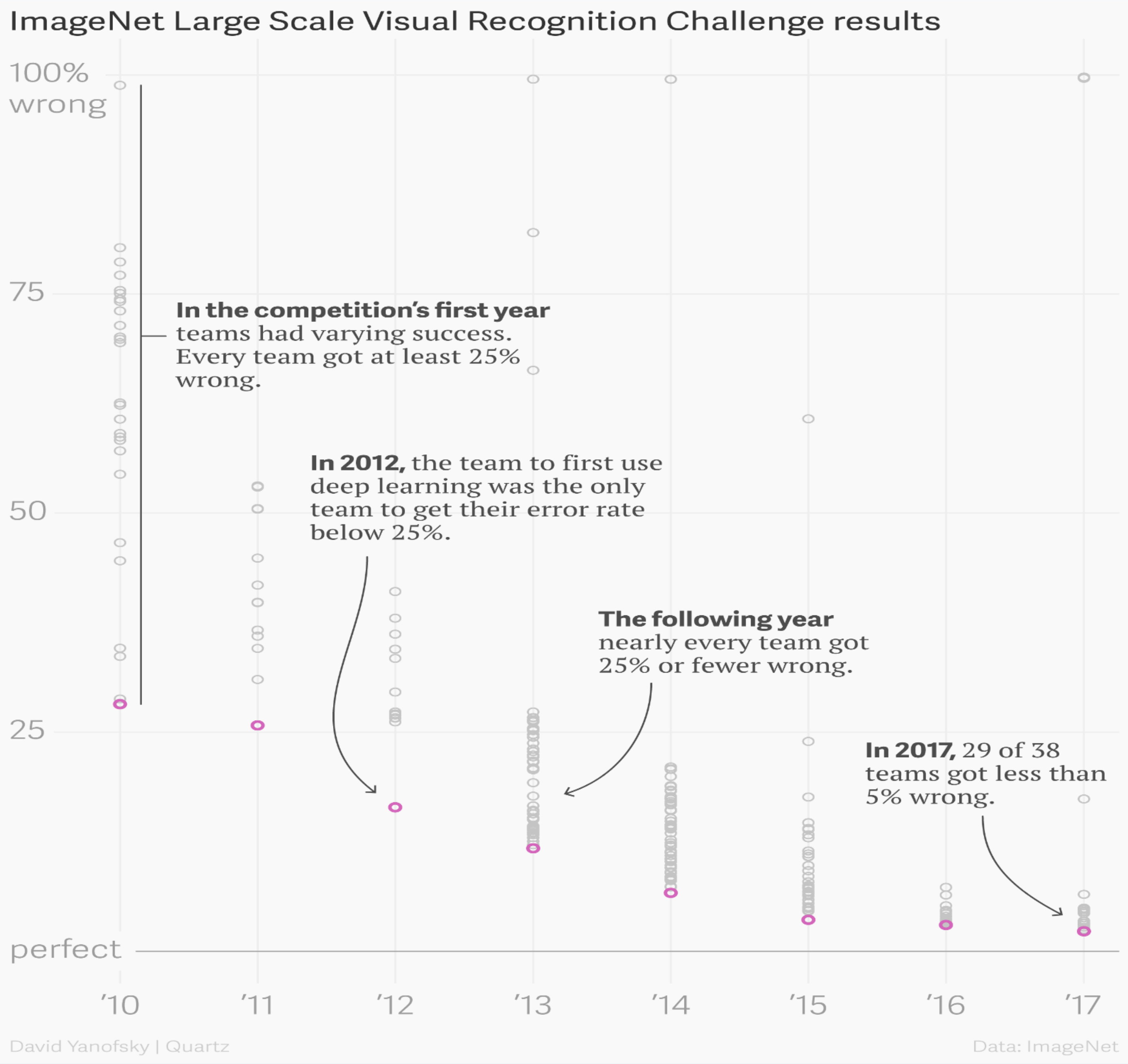


Figure 4: (Left) Eight ILSVRC-2010 test images and the five labels considered most probable by our model. The correct label is written under each image, and the probability assigned to the correct label is also shown with a red bar (if it happens to be in the top 5). (Right) Five ILSVRC-2010 test images in the first column. The remaining columns show the six training images that produce feature vectors in the last hidden layer with the smallest Euclidean distance from the feature vector for the test image. (Krizhevsky et al 2012)



Human top-5 classification error rate on the large scale ImageNet dataset has been reported to be 5.1%, whereas a state-of-the-art neural network achieves a top-5 error rate of 3.57% (Dodge 2017)

Figure 1. (a) Chihuahua and muffin, (b) Labradoodle and fried chicken

The Chichuahua and Muffin Problem. Togootoggtokh and Amartuvshin (2018) – used 20 layers

Why Did Machine Learning (ML) Explode?

1. Conceptual break through in algorithms – deep learning, boosting
2. The use of NVIDIA's GPU cards running CUDA – major speed gains
3. High quality free software: TensorFlow, XGBoost, Scikit-learn, etc.
4. Massive increase of machine readable datasets
5. Kaggle.com competitions have helped to popularize, and show what works

Now:

1. A GPU "arms race" is currently underway: NVIDIA, Google TPU, Intel Nervana, AMD, and many others; mainly in the USA and China
2. An "arms race" to set software standards: TensorFlow, Keras, PyTorch, Cognitive Toolkit, MXNet, Gluon, PaddlePaddle, etc.
3. So far the main players (Google, Microsoft, Facebook, Amazon, etc.) have been remarkably open: https://arxiv.org/, https://papers.nips.cc/, etc.

Part 3.
Since About 2015 Many Task for ML

- Siri, Alexa, etc.
- Read medical images about as well as a radiologist
- Convulsive seizure detection and activity tracking
- Various medical applications; genomics
- Modern cameras know to focus on a face
- Using drones and ML to identify crop disease areas, even before humans can tell
- Dominance at Go, Chess, Shogi, Texas Hold'em, Jeopardy
- Acting as a TA for a comp sci class (GA Tech) without the students realizing
- Part of training self-driving vehicles
- Robotics and drone controls
- >800 million hours of surveillance video collected daily
- Machine translation between languages is improving rapidly
- Google and Bing search results,
- Amazon prices (Bajari, Chernozhukov, Hortaçsu, Suzuki, 2018)
- Portfolio advice for investors
- Fraud detection systems
- High frequency trading systems
- AutoML, computer generated computer learning
- Generative adversarial nets produce much more realistic looking pictures
- … etc.

Expect to see some "Modern Luddite" reactions, but not much yet

Well Defined Inputs and Clear Outputs

- 2011 IBM's Watson machine wins Jeopardy game show
- January 2017 Libratus AI in a 120,000-hand heads-up no-limit Texas hold'em competition. Defeated four top human professionals.
  - The top ranked human, Dong Kim: "I didn't realize how good it was until today. I felt like I was playing against someone who was cheating, like it could see my cards. I'm not accusing it of cheating. It was just that good."
- December 2017, DeepMind's Alpha Zero algorithm (on high end hardware) learns to play chess, go, and shogi from scratch. Dominates the top fine-tuned computer programs at these games.
- January 2018, Microsoft and Alibaba AI programs exceed human standards on Stanford reading comprehension test for 1st time

Machine Solutions May Be Surprising

- Dec 2017, AlphaZero taught itself to play chess at superhuman levels and defeated Stockfish (a leading chess program). Also learned Go and Shogi well enough to beat the top programs

- The style of play is not like normal computer chess
  - AlphaZero “seems to understand positions much better than Stockfish”
  - Humans evaluate very few positions per second, Stockfish evaluates 70 million, AlphaZero evaluates 80,000. Time is going into deeper evaluations instead of more evaluations.

- “seeing the games I thought, well, I always wondered how it would be if a superior species landed on earth and showed us how they play chess. I feel now I know.”
  - GM Peter Heine Nielsen, the longtime second to World Chess Champion GM Magnus Carlsen

- ML solutions are likely to contain more such surprises

Where it is Starting to Work: Less Well Defined Settings

- 2016 “Jill Watson” is one of several TAs for a large undergraduate class at Georgia Tech who all communicated electronically. The students did not notice that, unlike the other TAs, she is not human. She is IBM’s program trained on a database of past student questions and answers for the course.

- 2010 to 2017 major improvements in Alexa (Amazon), Siri (Apple) and others, natural language communication with computers
- 2017 Volvo signs a deal to provide 24,000 self-driving cars to Uber

- 2017 Emotional algorithms?
  - Examples: https://www.soulmachines.com/, and https://www.affectiva.com/
  - Making the machines seem more natural to humans.

ML Designed ML Algorithms: Cloud AutoML

- AutoML, Machine Learning for Automated Algorithm Design, https://cloud.google.com/automl/
  - "Building and optimizing a deep neural network algorithm normally requires a detailed understanding of the underlying math and code, as well as extensive practice tweaking the parameters of algorithms to get things just right." …
  - "Google researchers have been testing the limits of automating AI for some time now. In 2016, one team showed that deep learning could itself be used to identify the best tweaks to a deep-learning system. Last year another group at the company used simulated natural selection to "evolve" an optimal network architecture." (MIT Technology Review, January 17, 2018)

Part 4.
Finance Practice is Moving

- "World's largest hedge fund [Bridgewater Associates] to replace managers with artificial intelligence"(The Guardian, December 2016)
- "The world's biggest money manager [BlackRock] on Tuesday announced that it would cut more than 40 jobs, replacing some of its human portfolio managers with artificially intelligent, computerized stock-trading algorithms. (Fortune, March 2017)
- "Goldman Sachs shows just how devastating automation can be to traders. In 2000, its U.S. cash equities trading desk in New York employed 600 traders. Today, that operation has two equity traders, with machines doing the rest." (Newsweek, March 2017)
- "Investment groups have more than quadrupled their number of "alternative data" analysts over the past five years, as asset managers scramble to unlock the potential of trading signals contained in website scrapes, language analysis, credit card purchases and satellite data." (Financial Times, January 29, 2018)
- "The world's biggest investment group, with $6.3tn of assets under management, is establishing a "BlackRock Lab for Artificial Intelligence" in Palo Alto, California, according to an internal memo seen by the Financial Times." (Financial Times, February 19, 2018)

Automating Routine Financial Decisions

- "fraud detection, capital optimisation, and portfolio management applications appear to be growing rapidly", Financial Stability Board (2017)
- Many day to day finance decisions can be automated since deep learning gives more accurate predictions of repayments
  - Credit granting: mortgages, loans, credit cards, etc.
  - What happens to the demand for loan officers?
- Giving financial portfolio advice
  - Increasingly automated and can be personalized: "Top 10 Robo Advisors Ranked: Find the Best Automated Online Investing Services" and "Best Robo-Advisors: 2018 Top Picks" (Vanguard, Schwab, and many new players)
  - What happens to the demand for human investment advisors?
- Answering customer questions, marketing to find new customers, etc.

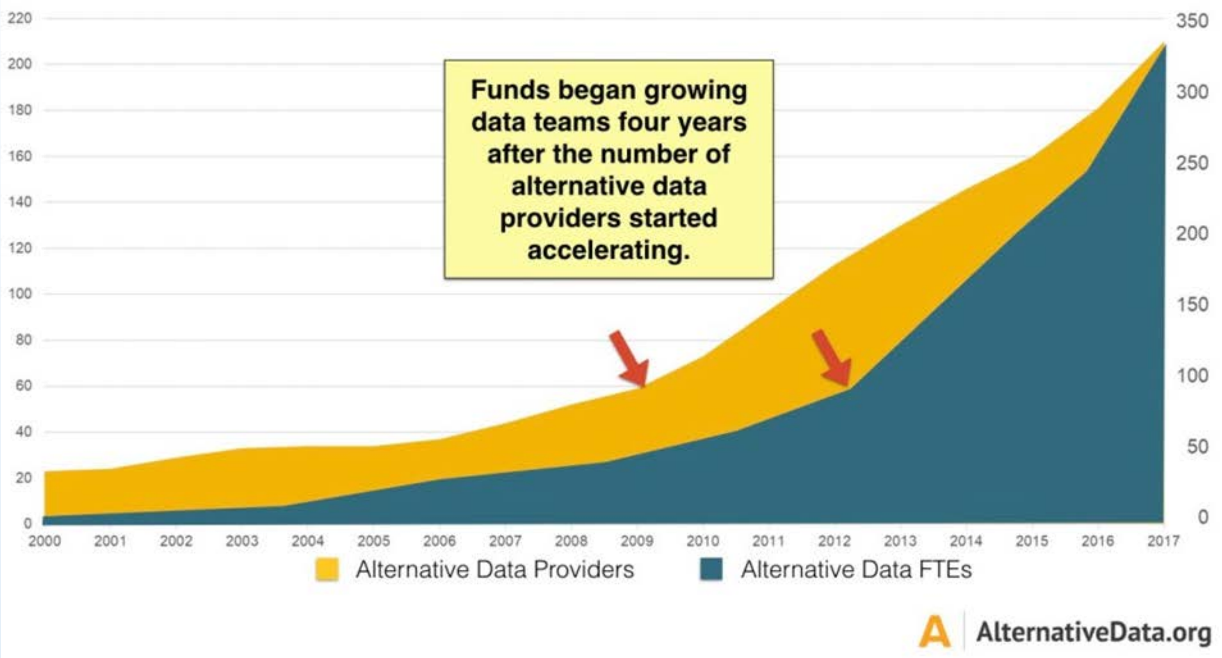
Regulators are Becoming Interested

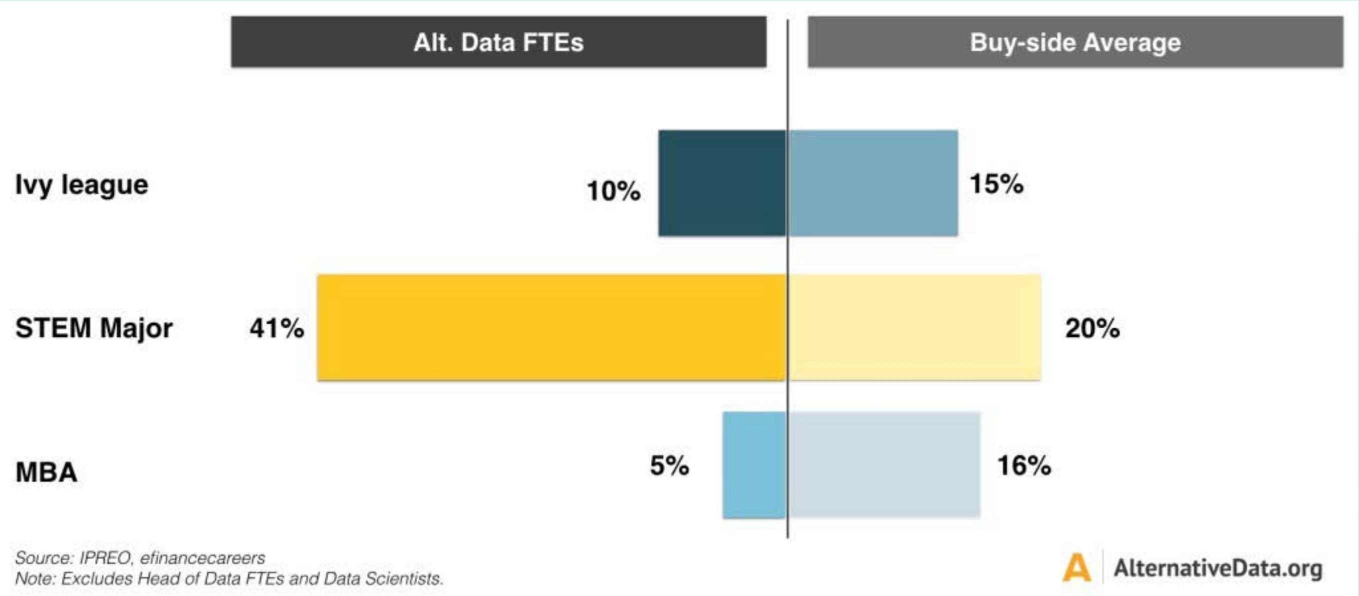
- It is claimed that this year half of business ethics violations will occur through the improper use of Big Data analytics (Guidotti 2018 et al)
- What to do if the algorithm inadvertently learns to discriminate by race, gender, etc? See Fuster et al (2018).
  - The deep learning algorithms are pretty good at finding robust patterns in the data
  - Suppose the law and ethics says some variable must be ignored, we may need special steps to force the algorithms to ignore those aspects
  - What if firms in other countries do not ignore them? They could cherry pick. Potentially a serious policy problem
- Wall (2017, Atlanta Fed)
- Financial Stability Board (2017)

Financial Stability Board (2017)

“Many of the models that result from the use of AI or machine learning techniques are difficult or impossible to interpret. The lack of interpretability may be overlooked in various situations, including, for example, if the model’s performance exceeds that of more interpretable models.”

“Yet the lack of interpretability will make it even more difficult to determine potential effects beyond the firms’ balance sheet, for example during a systemic shock. Notably, many AI and machine learning developed models are being ‘trained’ in a period of low volatility.”


Funds began growing data teams four years after the number of alternative data providers started accelerating.
220
200
180
160
140
120
100
80
60
40
20
0
350
300
250
200
150
100
50
0
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
Alternative Data Providers
Alternative Data FTEs
AlternativeData.org


Alt. Data FTEs
Buy-side Average
Ivy league
10%
15%
STEM Major
41%
20%
MBA
5%
16%
Source: IPREO, efinancecareers
Note: Excludes Head of Data FTEs and Data Scientists.
AlternativeData.org

Effects on Academic Finance

- **What can we do that the machines will not soon be ready to do?**

- Teaching
  - Using Neural nets to automate essay grading?
  - Automated TAs for large classes
  - What parts of teaching is better done by machine? By humans?
- Research
  - Some ML methods are starting to be used (Naïve Bayes, SVM, Lasso, Neural Nets)
  - Methods are being adapted to our needs (Causality in Random Forests, Deep Instrumental Variables, Double Selection Lasso, etc.)

- A likely drop in many mid-level finance jobs over the next decade.
- Jobs for our students? Fewer traditional jobs. Need changing skill sets.
  - Will there be an “ML Excel”?

In Some of My Own Work

- Antweiler and Frank (2004) “Is all that talk just noise? The information content of internet stock message boards” JF
- Frank and Sanati (forthcoming) “How Does the Stock Market Absorb Shocks?” JFE
- Frank and Yang (work in progress) “Does Finance Flow to High Productivity Firms?”

- In my experience, simple methods like Naïve Bayes, GXBoost, and Lasso work well

Part 5.
Conclusion: Change is Already Underway

- Academic finance research will benefit by using some ML methods
  - Traditionally, we mostly estimate parameters of a given model
  - Boosting, Lasso, Deep Learning methods often work remarkably well
  - We will also develop ML methods that match our needs (Athey, Chernozhukov, etc.)
- Teaching is already changing due to on-line. ML is likely to produce even bigger changes, but it will take time.

- Now is a time of tremendous experimentation – huge opportunities (and threats)

- Machine decisions may have a different character from human decisions
- Perhaps 15 or 20 years for finance to really adapt.
- Where will this leave us? I hope to talk it over with you in person – perhaps at the 2033 Midwest Finance Association meetings.

To Learn More

- Efron and Hasti (2016) “Computer Age Statistical Inference: Algorithms, Evidence and Data Science” Cambridge University Press, https://web.stanford.edu/~hastie/CASI/

- Goodfellow, Bengio, and Courville (2016) “Deep Learning” MIT Press. http://www.deeplearningbook.org/

- Competitions and discussion: https://www.kaggle.com/
- Popular research papers: http://www.arxiv-sanity.com/
- New data for investors: https://alternativedata.org/