Last Lecture

Advanced Machine Learning
Data Mining &
Artificial Intelligence
CSCI E-82
Fall 2018

Peter V. Henstock

Aggressive agenda tonight

- Eliud: perspective on NIPS
- HW6 discussion
- Brief intro to RL
  - Last two topic presentations on RL
  - Gerald Pho's application of RL to neurosci
- 1 final presentation
- Course wrap-up
  - Peter's Top 10 List for what's next in AI
  - Course survey
  - Conclusion

Types of Machine Learning

- Unsupervised
  - Clustering, PCA, etc.
  - Provide data without labels
- Supervised
  - Classification, regression
  - Provide data with labels
- Semi-Supervised
  - Classification but leverage clusters or distributions of unlabeled data
  - Provide data with some labels
- Reinforcement Learning
  - Choose best actions but find your own labels

Reinforcement Learning

1) Deterministic Search
2) Markov Decision Process
3) Reinforcement Learning

1) [Deterministic] Search

Example: “Grid World” Maze problem

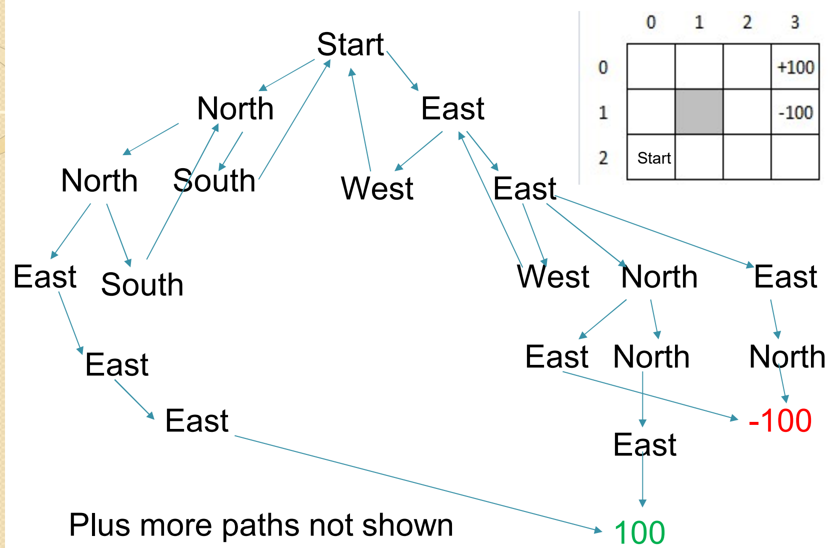
- Find the path to maximize reward
- Run multiple approaches = search
- Answer: shortest path & max reward

	0	1	2	3
0				+100
1				-100
2	Start			

- How would you solve this?
- What would a solution look like? #steps?

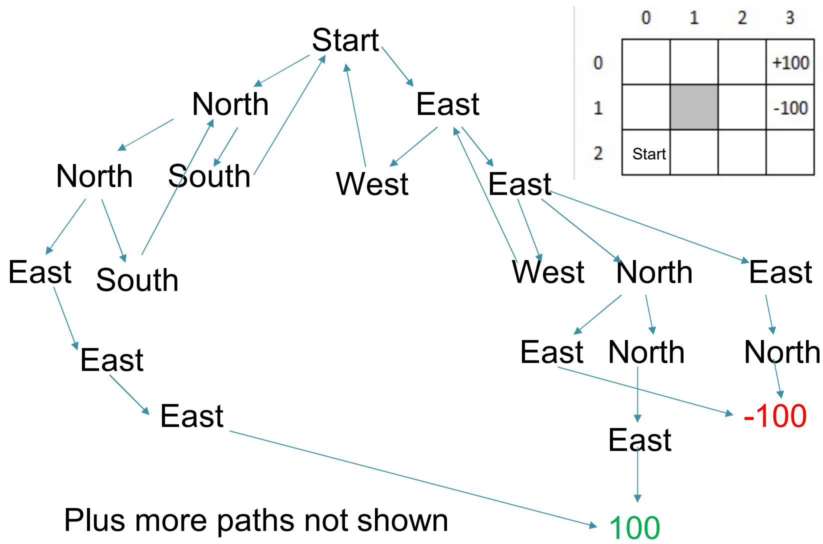
https://galweejit.wordpress.com/2010/12/16/ai-class-implementation-of-mdp-grid-world-from-week-5-unit-9/

Network or Decision Tree Idea



When do you get your reward?

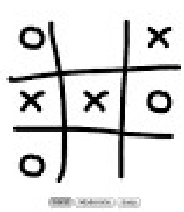
Network or Decision Tree Idea

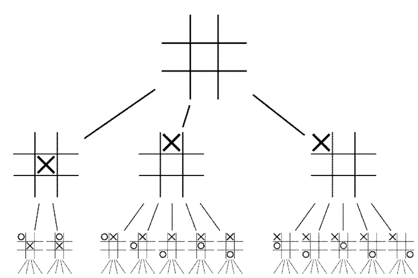


What is the best path?

How to play tic tac toe

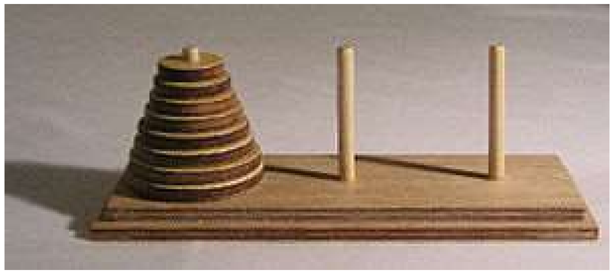
- AI search



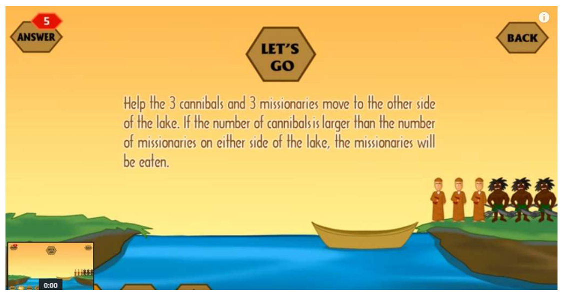


- Minimax solution:
  - Minimize the cost of your opponent maximizing their score
- What if you don't have a model of how opponent will play?

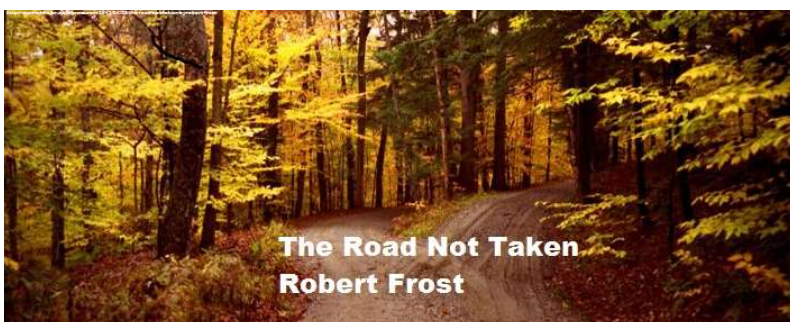
Tower of Hanoi



Cannibals & Missionaries



Two roads diverged in a yellow wood…



How do you determine which is the better direction: left or right?

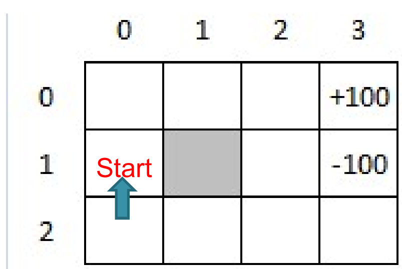
https://simeonpangsingwan.wordpress.com/college-writing-semester-2/poetry-analysis-of-favorite-poems/the-road-not-taken-by-robert-frost/

Solution for Traditional AI

- Define heuristic measures
  - Create a quantitative score that encourages the algorithm to make good choices
    - 2 in a row → 10
    - Blocking 3 in a row → 20
    - Block 3 in a row and getting 2 in a row 30
    - Middle edge → 5
    - Corner → 7
- Use search to find efficient algorithms
  - A* heuristic search
  - AO* heuristic search

[Deterministic] Search

Example: “Grid World” Maze problem
- Find the path to maximize reward
- Run multiple approaches = search
- Answer: shortest path & max reward



https://galweejit.wordpress.com/2010/12/16/ai-class-implementation-of-mdp-grid-world-from-week-5-unit-9/

- How would it change the problem?

[Deterministic] Search

Example: "Grid World" Maze problem
- Find the path to maximize reward
- Run multiple approaches = search
- Answer: shortest path & max reward

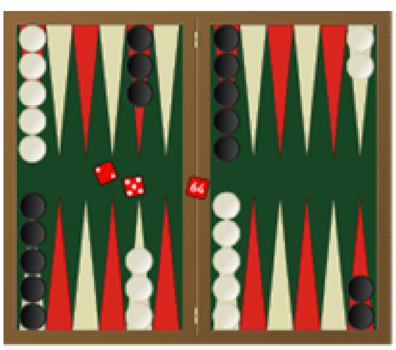
	0	1	2	3
0				+100
1	Start			-100
2				

https://galweejit.wordpress.com/2010/12/16/ai-class-implementation-of-mdp-grid-world-from-week-5-unit-9/

- Add -5 reward per step
- How would it change the problem?

Timing of the Rewards

- When do you get the rewards?




- Which is worth more:
  - 3% bond due in 5 years
  - 3% bond due in 25 years
- Discounted future rewards

2) Non-Deterministic Search

- Same input → different behaviors
- Standard: probabilistic algorithm

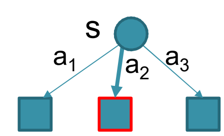
- Example: "Grid World" Maze problem
  - 80% of time action as expected (N,S,E,W)
  - 10% of time: 90° clockwise of target
  - 10% of time: 90° counter-clockwise of target

	0	1	2	3
0				+100
1				-100
2	Start			

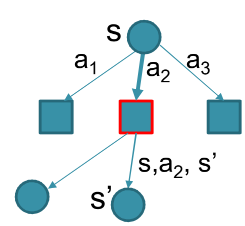
Process of switching states

Q-state is ■

State is ●



- Start in state s
- Choose an action a_2 that takes you to a "Q-state"
- Randomness shows transition to state s'



- (s,a_2,s') is a "move"

Markov Decision Process

- Framework using probabilities to optimize
- Take allowed actions from state s → s'
- Get a reward for taking certain actions

- Set of states $s \in S$
- Set of actions $a \in A$
- Reward function R(s, a, s')
  - Utility = Σ rewards
- Transition function T(s, a, s')
  - Model of the world or probability

- Why is it a Markov Decision Process?

Markov Decision Process

- Why is it a Markov Decision Process?
  - s' depends only on previous state s and action a and not all earlier states/actions

$$P(s' \mid s, a) =$$
$$= P(S_{t+1}=s_{t+1}' \mid S_t=s_t, A_t=a_t, S_{t-1}=s_{t-1}, A_{t-1}=a_{t-1}, S_{t-2}=s_{t-2}, A_{t-2}=a_{t-2}, \ldots)$$

- Because it's Markov…

$$= P(S_{t+1} = s_{t+1}' \mid S_t = s_t, A_t = a_t)$$

Solving MDP problems

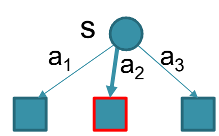
Several related problems as solutions

- Assess value of being in a given state based on future rewards
- Assess the future [discounted] rewards from taking a given action from a given state
- Finding optimal actions or policies

Recursive Optimization

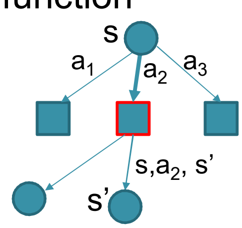
$V^*(s) = \max_a Q^*(s,a)$

- q-state is ■ state is ●



$Q^*(s,a) = \Sigma_{s'} T(s,a,s')[R(s,a,s') + \gamma V^*(s')]$

- Average over all children s'
- Prob of given s' = Transition function



Start off $V^*(s) = 0$ & iterate

$V^*(s) = \max_a \Sigma_{s'} T(s,a,s')[R(s,a,s') + \gamma V^*(s')]$

$V_{k+1}(s) = \max_a \Sigma_{s'} T(s,a,s')[R(s,a,s') + \gamma V_k(s')]$

Work backwards to compute V's

- Start at end and have 100 or -100

  White box = 0.0

	0	1	2	3
0	0	0	0	
1	0		0	
2	Start	0	0	0

- What if 1 step away, V*(s) =

	0	1	2	3
0	0	0		
1	0		0	
2	Start	0	0	0

- What if 3 step away, V*(s)

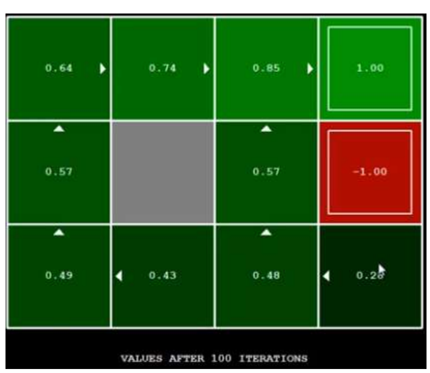
	0	1	2	3
0				
1	0			
2	Start	0		0

- Results will converge

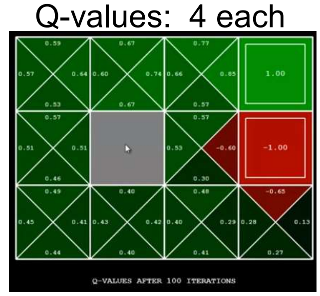
Grid World after 100 iterations

V*(s) shown in each square

- Expected discounted value of starting in square and acting optimally
- Arrows are optimal policy



Q-values: 4 each

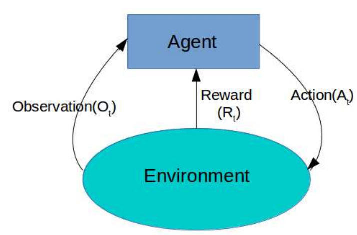


MDP ideas

- Started off with an MDP game rule with the stochastic process
  - Compute Value of each state
  - Compute policies of each state

- What if don't know what the transition probabilities are?
- What if don't know reward function?
  - We'd have to learn these
- How would you learn these?

Reinforcement Learning

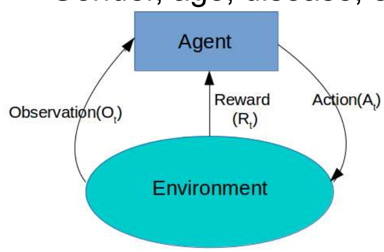


- Environment has states $s \in S$
- Agent makes actions $a \in A$
- Rewards $r \in R$ possibly for each action
- Markov assumption $P(s_{t+1} \mid s_t, a_t)$
  - Next state depends only on previous state and the action from that state to next
- Don't know T or R but do exist as MDP

Parts of Reinforcement Learning

- Observations
- Agent
- Action
- Feedback

Gender, age, disease, etc.



Web page, ad, treatments

- Example:
  - Show which ads to which users
    - Goal: maximize profit
  - Treat patients using different medicines
    - Maximize patient quality of life

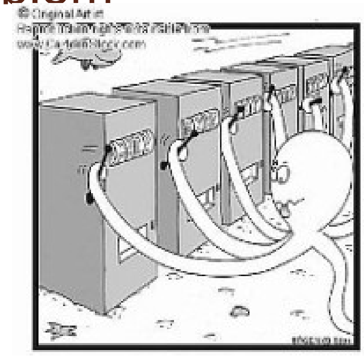
https://becominghuman.ai/components-of-an-rl-agent-and-its-application-on-snake-1b3b6c8e1de5

iRobot Roomba
Bedford, MA

- Goal = clean my house
- Explore: doesn't know how many rooms
- Exploit: has to clean the full space
- Constraint: limited power → recharge

N-Armed Bandit Problem

- Goal is maximize $
- 1000 actions
- Slot with N levers
- Maximize return



http://www.socsci.uci.edu/~szhang/pics/gamble.jpg

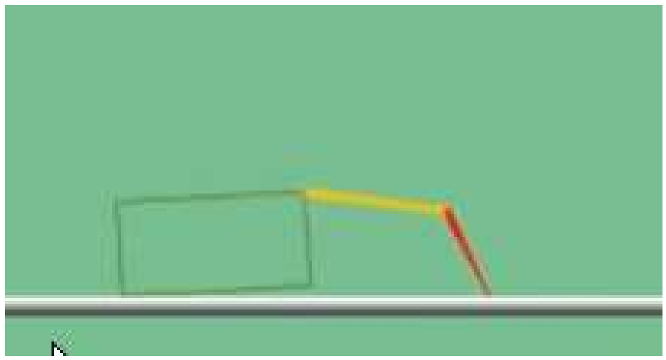
- Doctor has sick patients
- Doctor has multiple treatment options
- Reward is successful treatment or survival of patients

Concept of Reinforcement Learning

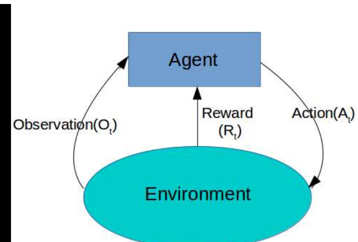
- Builds a model of its environment
  - May choose many bad paths in the process
  - Learns the transition state probabilities
  - Learns rewards
- Strategies in play (Q-learning)
  - Exploration (small probability)
  - Exploitation (large probability)
- Advantage:
  - Gathers own data and teaches itself

Graphical Robot

- How to train it to move the “shoulder” and “elbow” to move to the right?



Space Invaders

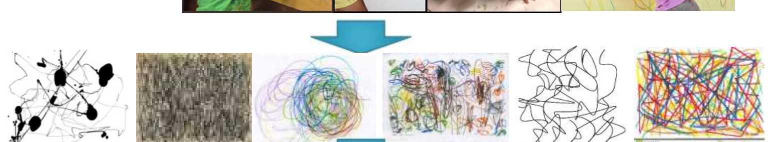


- How does the game map to RL?

~~Dave's~~ Peter's Upcoming AI/ML Trends

#10 Adversarial ML such as GANs

Forgers =
"Generators"



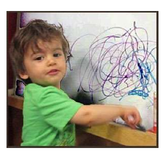
Inspector =
"Discriminators"

Jackson Pollock

One: Number 31, 1950

#10 Adversarial ML such as GANs

Forgers =
“Generators”



Inspector =
“Discriminators”

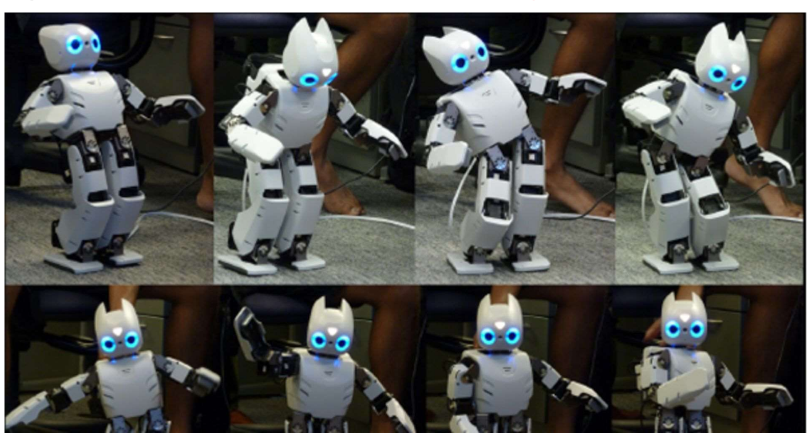
#10 Adversarial ML such as GANs

Forgers =
“Generators”

Inspector =
“Discriminators”

Peter’s Upcoming ML/DM/AI Trends

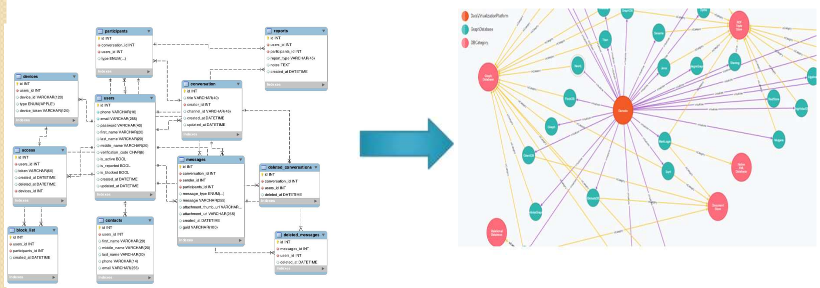
10) Further adversarial networks
9) Reinforcement Learning on the rise



https://www.cnbc.com/2015/12/04/a-baby-step-on-way-to-robots-learning-every-human-thing.html

Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models



Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP

Open-Source

- High quality open-source tools for machine learning & data mining
- Google, Facebook and universities release all their solutions
- AI vendors incorporate the latest tools
  - RapidMiner, H2O, DataIku, Weka
  - DataBricks, Hadoop, etc.
  - IBM

Who Owns the Intellectual Property?

- Small companies are not leading
- Google, Microsoft, Uber, Baidu, Facebook, Didi are leading
  - Hiring the top talent
  - Collaborating with top universities
  - Patenting the latest scalable technologies
- Difficult to maintain talent
- Difficult to compete against giants
- Multitude of problems to solve

Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP
6) New deep learning architectures 4 yrs

Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP
6) New deep learning architectures 4 yrs
5) Interpretability & Accuracy

Cassandra

- Apollo gave her ability to create prophecies in order to seduce her
- She refused
- Apollo cursed her so that no one would ever believe her
- ML/AI has this problem

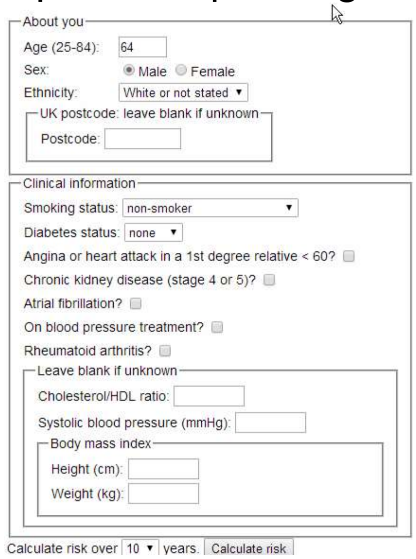
Cassandra by Evelyn De Morgan (1898, London);

Rule of 5

- Set of rules for predicting bioavailability of small molecules in humans from few thousand drug candidates
- Poor viability if:
  - Molecular Weight > 500
  - More than 5 H-bond donors
  - More than 10 H-bond acceptors
  - cLogP > 5
- What do you think of this model?

10 year Risk Coronary Heart Disease

- http://www.qrisk.org/



Emergency Room Protocol

- 10 year risk of coronary heart disease
- Ask patients 5 questions
- Count up scores and get a decision

- Two possibilities:
  - Model came from doctors
  - Model came from machine learning

- Which would you prefer?

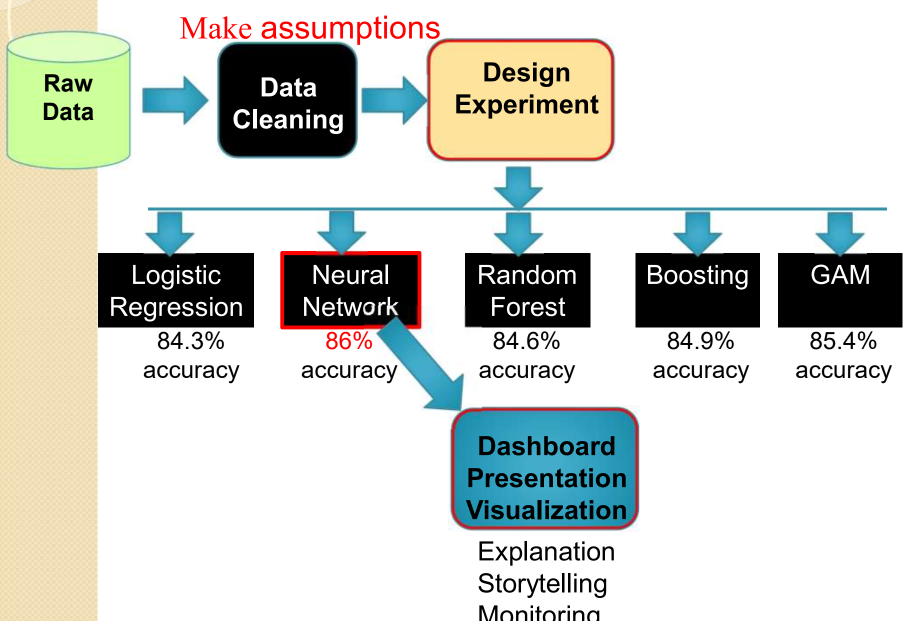
Case Study

Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission

Rich Caruana, Microsoft Research, rcaruana@microsoft.com
Yin Lou, LinkedIn Corporation, ylou@linkedin.com
Johannes Gehrke, Microsoft, johannes@microsoft.com
Paul Koch, Microsoft Research, paulkoch@microsoft.com
Marc Sturm, NewYork-Presbyterian Hospital, mas9161@nyp.org
Noémie Elhadad, Columbia University, noemie.elhadad@columbia.edu

- Goal: To predict probability of death for patients with pneumonia
  - Use machine learning to identify risk
  - Admit high-risk patients to hospital
  - Treat low-risk patients as outpatient
- 9847 patient training set, 4352 test set
- 46 features: age, gender, heart rate, BP...

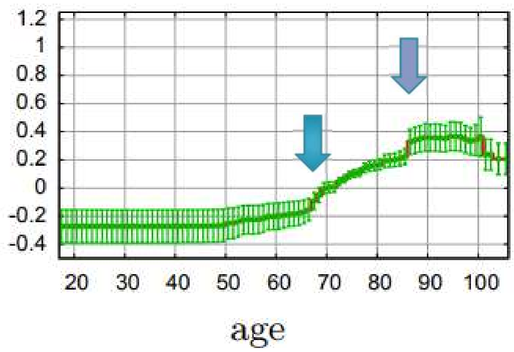
Data Science Tendency



Findings

Asthma patients → lower risk of death???
Chronic lung disease→lower risk of death???
Chest pain history → lower risk of death???

Risk by Age

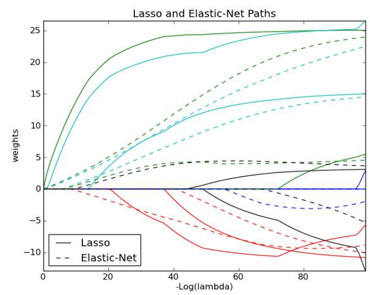


Accuracy & Interpretability

- https://www.nrcc.org/2015/11/06/brad-ashford-having-his-keystone-cake-and-eating-it-too/

Don't we have that already?

- Decision trees are most widely used mostly due to their interpretability
  - ID3
  - CART
  - C5.0
- Lasso
  - Regression method
  - Regularization



- http://jaquesgrobler.github.io/Online-Scikit-Learn-stat-tut/_images/plot_lasso_coordinate_descent_path_11.png

Bayesian Rule List

- http://web.mit.edu/rudin/www/LethamRuMcMa12.pdf
- Probabilistic Classifier like CART
- User can specify: - Number of rules
  - - Conditions per rule
  - - Priors of the output labels
- Steps
  - Mine data to come up with rules:
    - Apriori, FP Growth, decision tree, etc. if X → Y
  - Sample decision rule length~ Poisson(lambda)
  - Conditions ~ Poisson(eta)
  - Sample from rules that apply for conditions
  - Predict using first rule that applies from list

Opposite Trends

- Deep Learning
  - State of the art algorithms
  - Training data added up to the limits of GPU and training time
  - Better decisions that no one can understand
- Trend toward Simplicity
  - Rule of 5 (Lipinski)
  - Emergency Room Protocol: Risk of heart disease
    - Doctors: Ask 5 questions and count score
    - ML/AI: Ask questions and enter values to compute a correct interpretation

Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP
6) New deep learning architectures 4 yrs
5) Interpretability & Accuracy
4) Privacy, bias, and legal aspects

Data

- Currently every web site tracks us
  - How much longer do we allow this?
  - Who keeps our medical records? Blockchain
- Are our results biased?
  - Data collection methods are biased
  - How do we acknowledge, leverage & fix?
- Legal aspects of AI
  - If robots replace us, will they pay taxes?
  - Technology is not limiting self-driving cars
  - What does war look like with AI? Preventable?

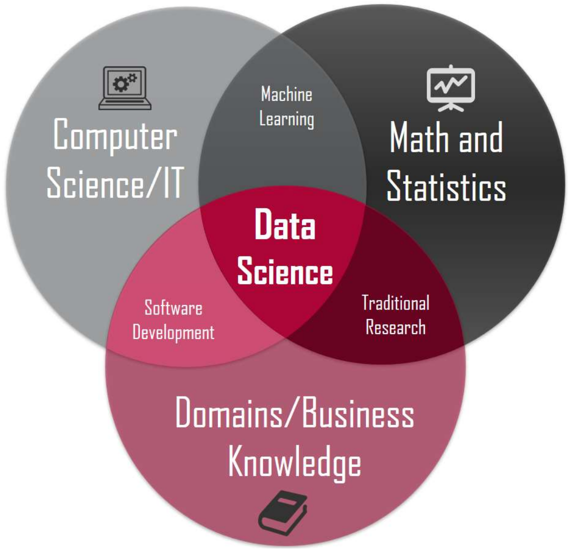
Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP
6) New deep learning architectures 4 yrs
5) Interpretability & Accuracy
4) Privacy, bias, and legal aspects
3) Shortage of AI talent in US

October 2012 HBR

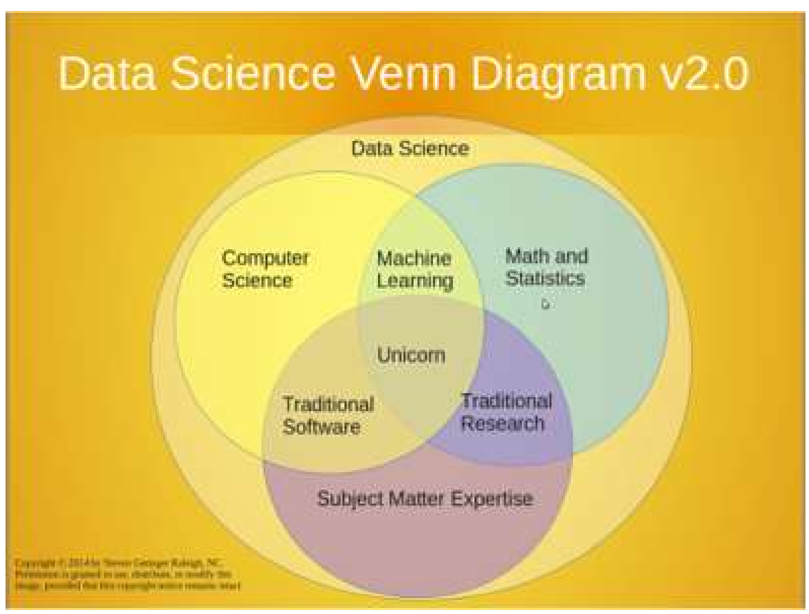
Yeah Baby!

Skill Landscape for Data Science & AI



https://towardsdatascience.com/introduction-to-statistics-e9d72d818745

"Unicorn"



- http://insidebigdata.com/2014/01/09/data-science-found-unicorn-yet/

©2018 Peter V. Henstock

Strong Shift of AI Talent to Tech Industry

- “Tech giants are paying huge salaries for scarce AI Talent” -- NYT October 22, 2017
- “Carnegie Mellon reels after Uber lures away researchers” -- WSJ November 21, 2018

Scarcity of Talent

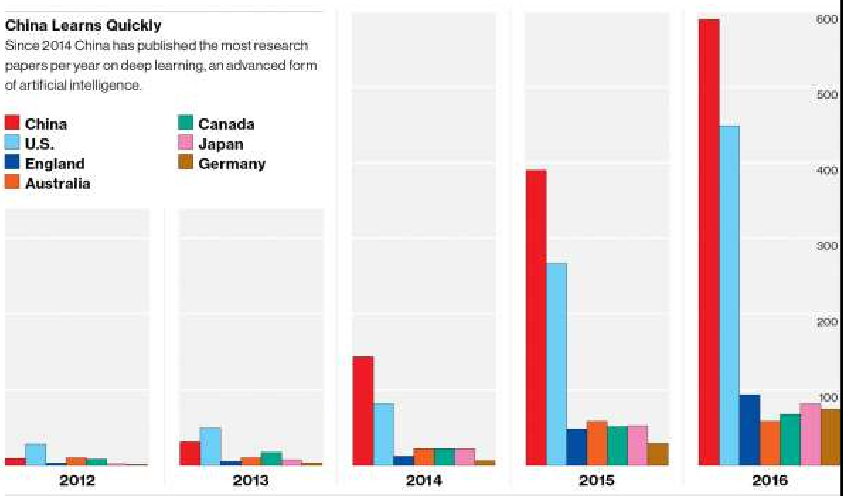
80% of machine learning engineers with PhDs are scooped up by Google and Facebook, especially if they have a patent. Forbes post: Gal Almog Feb 9, 2018

LinkedIn Jobs report:

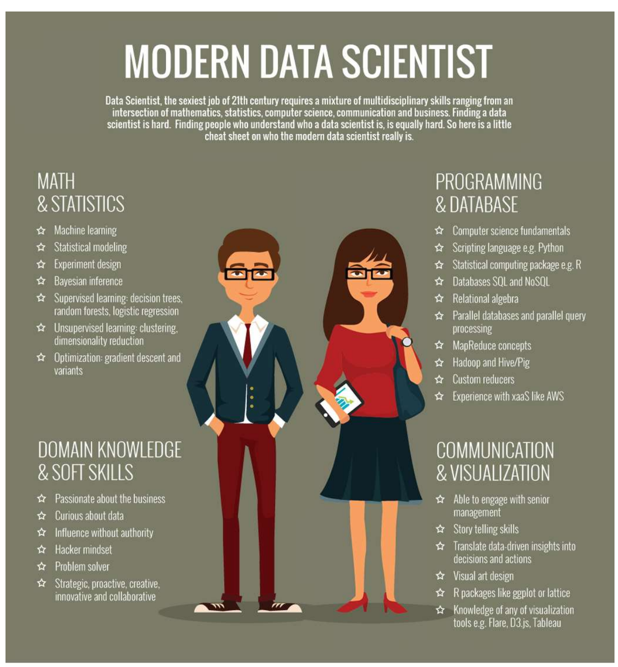
6.5x growth in data science jobs from 2012-2017

https://www.kdnuggets.com/2018/09/how-many-data-scientists-are-there.html

Deep Learning Papers by Country



https://www.technologyreview.com/s/608112/who-is-winning-the-ai-race/



MarketingDistillery.com is a group of practitioners in the area of e-commerce marketing. Our fields of expertise include: marketing strategy and optimization: customer tracking and on-site analytics: predictive analytics and econometrics: data warehousing and big data systems: marketing channel insights in Paid Search, SEO, Social, CRM and brand.

Needed Skills

- What do you do when it doesn't work?
  - Machine learning
- Is machine learning the right approach?
  - Statistics can often provide a better method
- How do you avoid resolving it again?
  - Software engineering
  - Shortage of SWE who can do data science
    - Holding the field back
- Business knowledge

Current need: data science

- Need in 2020 will be machine learning
- Data science algorithms have become quite accessible to everyone
  - Available on Oracle, python, R, Java, etc.
  - Many can push data through algorithms
- Competitive advantage of ML
  - How to go beyond the standard algorithms
  - What to do when the results are OK
  - How to achieve superior results

Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP
6) New deep learning architectures 4 yrs
5) Interpretability & Accuracy
4) Privacy, bias, and legal aspects
3) Shortage of AI talent
2) Integration of approaches

Pedro Domingos: U Washington

5 Tribes of Machine Learning (ACM)

- ttp://event.on24.com/eventRegistration/console/EventConsoleNG.jsp?uimode=nextgeneration&eventid=1087879&sessionid=1&username=&partnerref=&format=fhaudio&mobile=false&flashsupportedmobiledevice=false&helpcenter=false&key=DDB691DE07A8B2C54CCDC9CECDB98D3F&text_language_id=en&playerwidth=1000&playerheight=650&overwritelobby=y&eventuserid=129999792&contenttype=A&mediametricsessionid=103917850&mediametricid=1605789&usercd=129999792&mode=launch#

5 Tribes of Machine Learning

1) Symbolists
2) Connectionists
3) Evolutionaries
4) Bayesians
5) Analogizers

5 Tribes of Machine Learning

Tribe	Origins	Master Algorithm
Symbolists	Logic, philosophy	Inverse deduction
Connectionists	Neuroscience	Backpropagation
Evolutionaries	Evolutionary biology	Genetic programming
Bayesians	Statistics	Probabilistic inference
Analogizers	Psychology	Kernel machines

5 Tribes of Machine Learning

1) Symbolists: Use logic
   - Decision trees
2) Connectionists: model the brain
   - Neural networks
3) Evolutionaries: Nature's optimizations
   - Genetic algorithms/programming
4) Bayesians: Statistical inference
   - Naïve Bayes
   - Bayesian networks
   - Causal reasoning
5) Analogizers: Kernel machines
   - SVM as well as recommender systems

5 tribes of machine learning

- So what?

- Currently each tribe is doing its own research independently
- Opportunity is for tribes to converge

State of the Art (Domingos)

- Representation
  - Probabilistic logic
    - First order logic ←→ graphical models
    - Markov logic networks attach weights to logic
  - Weighted formula: distribution over states
- Evaluation
  - Posterior probability
  - User-defined objective function
- Optimization
  - Formula discovery (genetic programming)
  - Weighted learning (backpropagation)

Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP
6) New deep learning architectures 4 yrs
5) Interpretability & Accuracy
4) Privacy, bias, and legal aspects
3) Shortage of AI talent
2) Integration of approaches
1) Auto-ML

Peter's Upcoming ML/DM/AI Trends

10) Further adversarial networks
9) Reinforcement Learning on the rise
8) Graphs & more flexible data models
7) Push for both open-source & IP
6) New deep learning architectures 5 yrs
5) Interpretability & Accuracy
4) Privacy, bias, and legal aspects
3) Shortage of AI talent
2) Integration of approaches
1) Auto-ML
0) AI everywhere

IBM Commercials

- Watson predicts the weather
- Watson helps make our Olympians
- Watson cures cancer
- Watson solves education problems
- Watson learned to read and we sent it to medical school
- Watson designed this dress

- Watson raised the awareness of what AI can bring to the overall community

What area is next for AI?

- AI for data-driven decision making
  - When are you most efficient? What to do?
  - Scientific directions
  - Financial decisions
  - Political decisions
  - Legal decisions
  - Personal automated data-driven decision making

What area is next for AI?

- Your opinions?

What area is next for AI?

- Your opinions?

- Decision making
  - When are you most efficiencies? What to do?
  - Scientific directions
  - Financial decisions
  - Political decisions
  - Legal decisions
  - Personal automated data-driven decision making
  - Where does this leave your boss?

Currently Data Science is hot

- Standard functionality
  - Can you move data and call functions?
  - Can you interpret the results?
- Better than nothing, but high school skill

- Can AI replace data science? AutoML
  - Data Robot, H2O.ai feature eng., Auto-sklearn

- Data science vs. Machine learning
  - "Machine learning is part of data science"
  - Sure, try to do data science w/o it

Course in Review

How do you measure a course?

- 600 email threads 850 emails
- Piazza:
  - 1057 total contributions
  - 221 instructor responses; 28 by students
  - 12 min. avg. response time
- 240 homework/exams/projects
- 37 paper summaries
- 17 student presentations + code sets
- 1866 lecture slides
- 12 quality section notebooks

https://www.bionicturtle.com/forum/threads/good-luck-on-the-exam-tomorrow.9608/

Top 10 Data Mining Algorithms

- http://www.cs.uvm.edu/~icdm/algorithms/index.shtml
- Nice review paper: http://www.cs.uvm.edu/~icdm/algorithms/10Algorithms-08.pdf
- 2006 IEEE Data Mining Conference

1) C4.5
2) K-Means
3) Support Vector Machines
4) APRIORI
5) Expectation Maximization
6) PageRank
7) AdaBoost
8) K-Nearest Neighbors
9) Naïve Bayes
10) Classification and Regression Tree (CART)

Topics Covered Us vs. Them

Course Areas	Udacity supervised	Udacity unsupervised	Udacity reinforcement	Udacity ML	Johns Hopk coursera practical JH	Coursera Cluster Analysis	Coursera Pattern Discovery	Coursera Text Retrieval	Coursera Standford ML	Coursera Toronto NeuralNe	Coursera Stanford Min Mass	Coursera Uwash ML	Harvard CSCI E-181	Harvard DataSci
decision trees	x			x	x						x	x		x
regression	x			x	x				x				x	
classification	x											x		
neural networks	x								x	x			x	
instance based learning	x													
ensemble b&b	x											x		x
kernel methods and svm	x			x					x		x		x	x
computational learning theory	x											x		
VC dimensions	x											x		
Bayesian learning	x													x
Bayesian inference	x											x		
Naïve bayes	x											x		
clustering		x		x		x			x		x	x	x	
feature scaling		x		x	x									
text learning				x								x		
feature selection		x		x										
pca				x					x	x	x		x	
evaluation metrics				x	x									
markov decision			x										x	
reinforcement			x										x	
game theory			x											
randomized optimization									x					
information theory		x										x		
boosting					x									
regularized					x								x	
EM						x						x	x	
sequential patterns							x							
linear algebra									x	x				x
logistic regression									x					
outliers									x					
convolution networks										x				
boltzman										x				
pagerank											x			
lSH											x			
nearest neighbor											x	x		
recommender systems											x			
HMM													x	
deep learning												x		x
rule-based												x		

Your achievements

- PCA, eigenvectors, t-SNE, visualization
- Word2vec, topics, trend detection
- Regression & time series modeling
- Clustering and classification
- Deep learning with CNNs
- Text mining
- Recommender systems
- Network analysis
- Outlier analysis

Great Topic Presentations + Code

- Ensembles, XGBoost, GANs, SMOTE
- Density clustering
- Sentiment analysis
- ICA & dynamic PCA, UMAP
- Local outlier factor
- Time series change, clusters & MVA
- Network clustering
- Collaborative Filtering
- REINFORCE and Deep RL

Machine Learning Journal Sept.

- ALR accelerated logistic regression
- Fast and scale Lasso
- Factorizing LambdaMart for cold start recommendations
- Topological insight into deep learning
- Adaptive trajectory analysis of … clustering
- On need for structure modeling in sequence prediction
- Cost sensitive boosting algorithms
- …

KDD 2017 Conference Sections

- Networks & graphs
- Intelligent systems & data science
- Methodology
- Novel applications
- Representations
- Matrices
- Clustering
- Recommendations
- Supervised Learning
- Humans & crowds
- Medical data
- Deep learning
- Anomaly Detection
- Kernels and Sketches
- Temporal Analysis

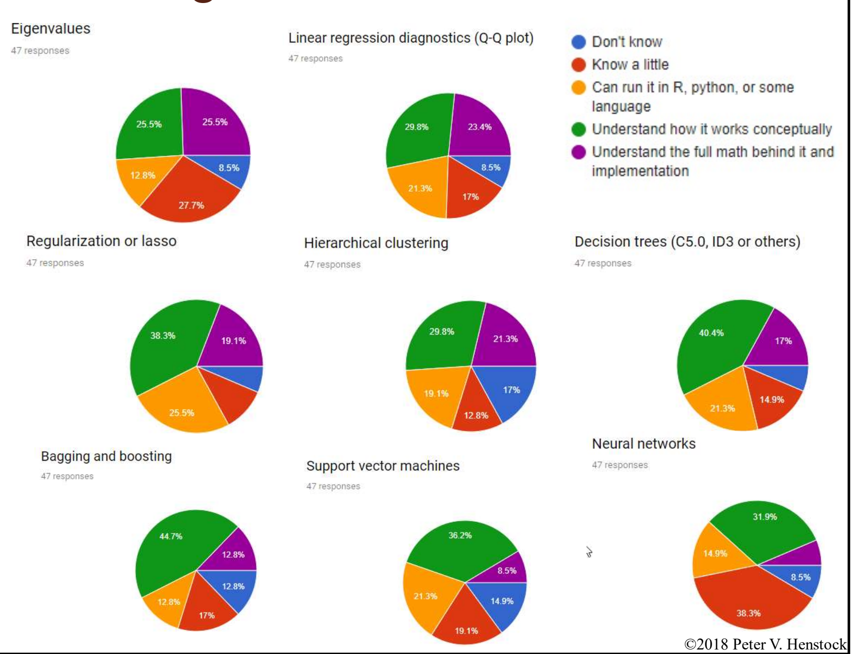
Not just the ML/DM/AI areas

- Bioinformatics Journal
  - DepQA: Improving estimation of single protein model quality with deep belief networks
- JCIM (Computational Chemistry)
  - Boosting docking-based virtual screening with deep learning
- Agriculture
  - Combo of fuzzy logic and analytical hierarchical process techniques to assess K saturation in soils….
- Journal of Politics
  - Strategic retirements of elected & appointed justices: Hazards model approach
- Economics
  - Tariffs, Trade & Productivity, Quantitative Evaluation of Heterogeneous Firm Models

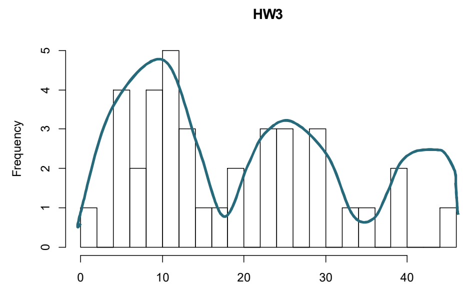
Goals

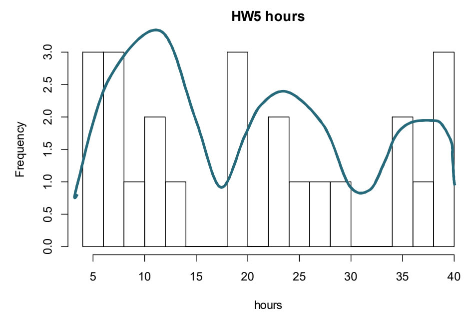
- Learn core ML/DM techniques at a moderate level so you can quickly go into more depth later
  - Statistics & time series, DTW
  - Unsupervised: Clustering, PCA, visualization
  - Supervised: classifiers, deep learning
  - Applications: text, freq. pattern, recommenders, networks, outliers, images, time series
  - Common methods: matrix, eigenvalue, optimization, entropy, EM
- Gain hands-on XP with state of art tools & applications
  - Used main python libraries scikit-learn & others
  - Deep learning, text mining, img. proc., classifiers, Kaggle, etc.
  - Keras, XGBoost, Scikit-Learn, AWS, Dask
- Leverage resources:
  - Present on recent research & read papers
  - Section/Exploratory notebooks = resources for you
- Apply skills: final project

Challenge of this 2018 course

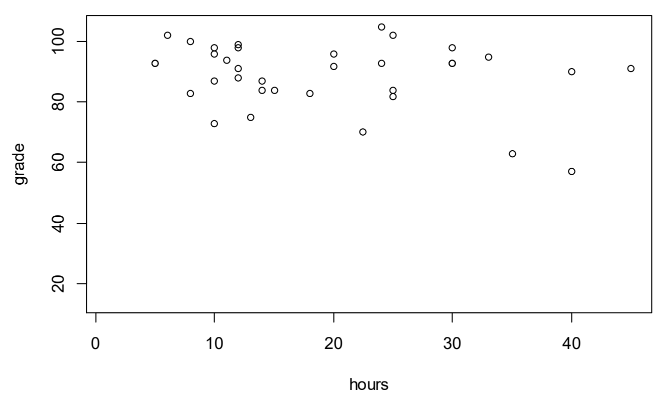


Histogram of Homework Hours





Everybody learned



It took more time for some, but everyone put in the time to succeed

Conclusion

Personal Ideas

- Differentiate your skill set
  - Difficult to be the best ML expert
  - Difficult to be the best programmer
  - Difficult to be the best anything
  - Combinations have massive opportunities
- Data science is the manifestation of this

3 Types of ML/DM/AI jobs

- Solver of all problems
  - Technical wizard to solve any problem
  - Solve different problems each week
  - Sometimes exploit knowledge from one to work on another
  - Frequently working alone
- Solver of one problem
  - Improve “revenue” or “predictions”
  - Basis of the business model
  - Every increment → $
  - Frequently teams of people
- Implementer

Recommendations

- Become more fluent in python or R
- Choose 1-2 favorite classifiers
- Choose 1-2 clustering methods
- Choose 1-2 higher dimensionality visualization approaches

- Learn the space where you apply them
- Be able to interpret results

Problem Solutions

- Visualize your data
- Try your favorite methods
- Start with solutions out of the box

- Identify a metric to assess results
  - Try multiple solutions
  - Adapt solutions based on:
    - Knowledge of the problem (visualize issues)
    - Research papers
  - Watch out for the bias-variance trade-off
- Communicate well: needs, results and insight to domain experts (not ML folks)

What are good next steps?

- ML & Stat: core of most non-obvious analyses
  - Use and maintain these skills in your toolbox

- Need more experience applying the tools
- Need a wider range of problems to push your envelopes into new areas

- Take on new project at work
- Let's research a problem together for a paper
- Work in the field: 6.4% raises in 2017 (Forbes)

Want more skills?

- Statistics & math (linear algebra)
- Graph analysis (social networks)
- Visualization
- Databases
- Text mining/linguistics/NLP
- Big Data
- Better coding: scripting & software eng.

- Don't forget domain expertise

Thanks to our Awesome TAs

- Dave & Rashmi

Thanks to our Awesome TAs

- Created fabulous notebooks each week
- Assembled examples of all the python concepts in first weeks
- Provided hands-on examples of how to convert lecture concepts into python
- Dove into sometime unfamiliar areas like Deep Learning, XGBoost, Spark
  - Dave's Deep Learning lecture
- Graded and provided great feedback

- Hope to offer this class next Fall
- Need TAs in the future
- Teaching will increase your knowledge
- Also might need TA for Spring Software Engineering Capstone if you are an experienced developer

Thanks to You

- Great group of people

- Challenged our ideas
- Presented new concepts

- Worked *very* hard all semester
- Contributed professionally
- Shared feedback to help us

Where are you now?

- Machine learning is core knowledge
  - Have exposure and moderate understanding across all the main sub-fields of the field
  - Can easily go deeper for a particular task
- Data science is the implementation

- You have more knowledge than most "data scientists" working in the field

  +

- You have proven you can implement a wide range of solutions in many domains

Final thoughts…

- I’d like to keep in touch with you
- Let me know if I can do anything for you now or in the future
- Go make the world better….

https://www.youtube.com/watch?v=P4LhWSN3YSw: