Raytheon

Developing Effective Technical Training

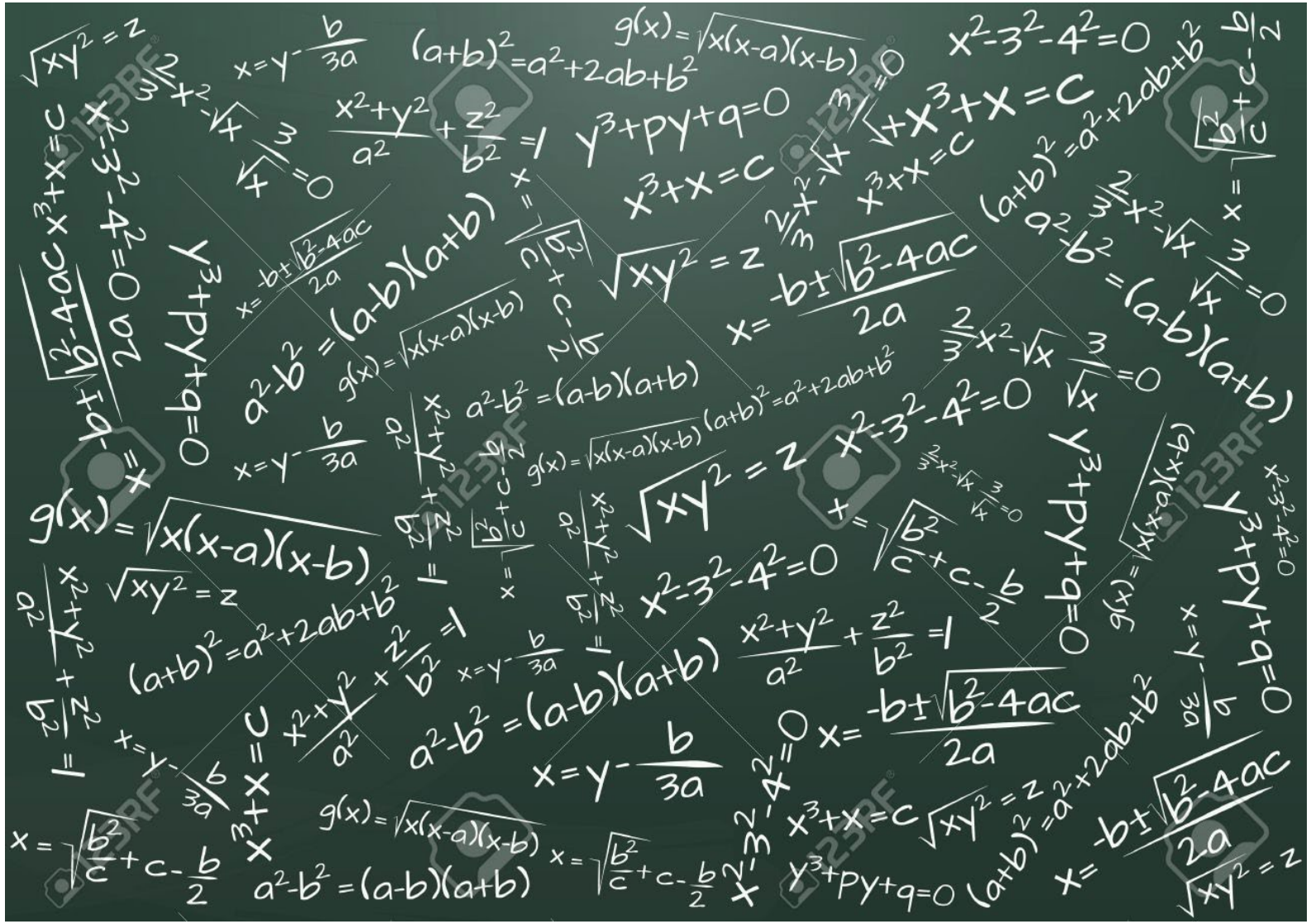
Tod Newman

Engineering Fellow, Raytheon Missile Systems
Lead, Raytheon Center for AI and Machine Learning

Developing Effective Technical Training

Raytheon

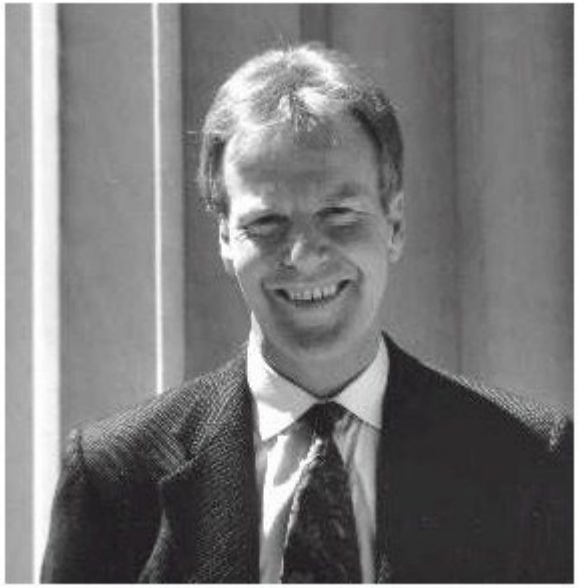
- What am I going to talk with you about today?



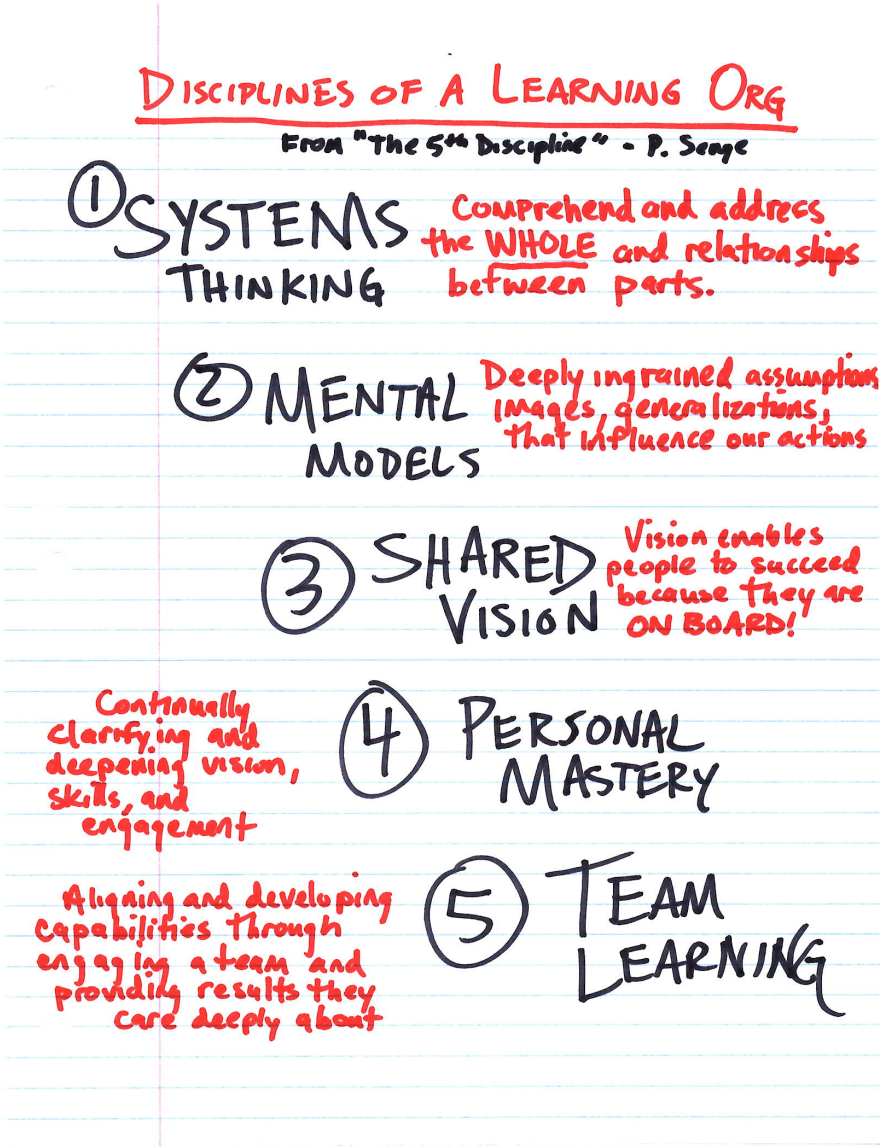
(maybe later)

Instead…

- Here's a subject that I think you can leverage to become more valuable if you land in a technical position in industry…
- Setting the stage:
  - Organizations have a habit of settling in to a routine and losing traction in the marketplace
  - Peter Senge of MIT ("The Fifth Discipline", 1990) has proposed the construct of a learning organization where:
    - "people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning to see the whole together"
    - These companies succeed over all others because in periods of rapid change, they can learn faster and adapt better.



Enabling the Learning Organization



- Much training exists already.
- Some is from people who's context was a past paradigm
- Some comes from vendors and was targeted to be general (and therefore profitable)!
- Most is mandatory and of little use.
- However, for the technologist who can build material that is successful at building thinkers, the sky is the limit

Why Should You Be Interested?

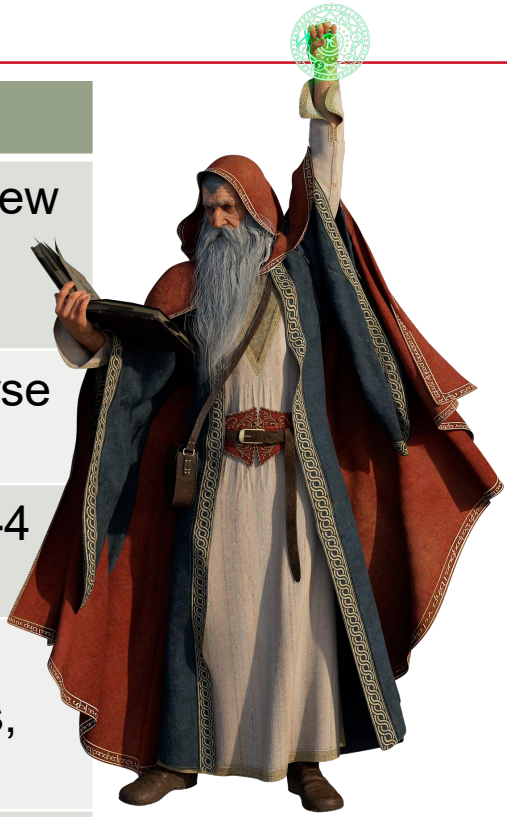
- In industry, to advance to senior technical roles, you need to understand:
  - Developing alignment around new tech and scientific developments
  - Maturation of new technologies
  - Preservation of niche science
  - Development of a talent pipeline – How to replace yourself!
- What fights against all this goodness?
  - Culture resists change
  - Organizational lack of awareness and focus
  - Buckets of money in the wrong places
  - Poor talent acquisition strategy

More on the Challenges

- Organizations struggle with Senge's Five Disciplines
  - Inability to see the whole destabilizes integration of new ideas, techniques, or business strategies
  - Existing staff's mental models are aligned around day-to-day sausage-making rather than improvements.
  - Shared vision is driven in areas of greatest risk aversion, not in areas of big opportunities
  - Personal mastery is in the context of short-term business goals, not long-term organizational health
  - Developing and empowering long-term teams opposes short-term staffing and sales goals.

Levels of Learning and Hollywood

Bloom Taxonomy	Target Learners	Hollywood Version	Training Formats
Remember	General Interest and Organizational Leaders	Short trailer (<1 min)	Introductory Overview (~15 min)
Understand	Functional Leaders	Full trailer (2-4 min)	Full Overview Course (90-180 min)
Apply	Practitioners	Movie	Training Course (1-4 days)
Analyze	Experts	Book	In depth courses (~semester), books, research
Evaluate	Mentors	Making of, interviews, expert and casual commentary	Certification programs, degrees
Create	Thought Leaders	Concepting, Scriptwriting	Original work, developing new courses, PhD



Learning wizardry: from Harry Potter movie trailer to PhD

Ideas on How to Build Transformational Training

Raytheon

- What is transformational about training?
  - Intentionally designed training regimes can have incredible impact on an organization
- I believe mapping training to the Disciplines of a Learning Organization is effective
- If you are fortunate enough to convince executives that you need to build a training program for your discipline, the following slides have some tips.

- Systems Thinking: The training program must be designed as a system of interconnected parts that fulfils a strategic purpose
- Mental Models: The training developer HAS to attempt to understand and account for the culture, shared experience, or the place in life of the targeted audience.
- Shared Vision: John Kotter's change model starts off with the developing of a vision of urgency. An effective training program does the same.
- Personal Mastery: More than just developing skills, advancing personal mastery involves deeper understanding and applications.
- Team Learning: Teams learn effectively together and leveraging teaming well can result in students internalizing training.

Techniques to Optimize your Technical Training Programs

Raytheon

- Develop Concepts with Personal Relevance
  - Personal Relevance: A learning environment characteristic emphasizing concrete, personally-relevant experiences to aid the learner in constructing individual meaning (Ke & Kwak, 2013).
  - Think about Cognitive-optimized strategies
    - Targeted Repetition
    - Building concepts from summary to details
    - Active Learning (workable examples, challenges)
    - Metacognition – students ID their top cognitive learning approaches
  - Develop Systems of Engagement
    - Student teaching of concepts
    - Analysis of Problems

Online learning across ethnicity and age: A study on learning interaction participation, perception, and learning satisfaction, Fengfeng Ke, Dean Kwak, 2013 (https://doi.org/10.1016/j.compedu.2012.09.003)

Techniques to Optimize your Technical Training Programs

Raytheon

- Teach Cultural Insertion
  - Primarily talking about corporate culture here, but could be any relevant culture to the business
  - Culture is reflected in the stories of goats and heroes and how they're rewarded
    - So stories of successes and failures ought to be part of your technical training strategy
    - Teaching change management (i.e., John Kotter model) in the context of your technical domain could also be beneficial
    - Make it clear that success is more than an elegant algorithm but is rather the work to both design and deploy that algorithm into 'production'.

Techniques to Optimize your Technical Training Programs

Raytheon

- Ensure Technical Concepts are Taught Within the Business Context
  - There is a place for teaching basic theory of your field, but better results can be had for teaching how the subject is deployed in your business.
    - This information is generally much more highly sought after by those with less experience than the fundamentals
      - Students can catch up on the fundamentals through short courses or the web, but can only get application details from you.
    - Again, there is a place here for detailed real-world examples and stories

Techniques to Optimize your Technical Training Programs

Raytheon

- Incentivize Practice
  - A person with an advanced degree will usually learn best in your business through carefully planned opportunities to practice the discipline
  - Perhaps the most important aspect of your training program will be the hands-on portion.
    - A system of working examples that cover many business eventualities will be critical for success.
      - Spend more time on this part of your training than the slick Powerpoint slides!
    - A framework that enables your students to work the examples using their assets and resources will incentivize them to practice on their own.
    - A contest (i.e., Kaggle) will often get students into the tools and examples faster than anything else.

Techniques to Optimize your Technical Training Programs

Raytheon

- Follow-up with Community Building
  - For true success in your training program, it should be structured as a select community
    - Certifications or qualifications could be designed for attainment through the program
    - Systems to connect graduates after the course that also aid collaboration will be very important
      - It is hard to keep content flowing to the community, so you may want to delegate this activity
      - Schedule community events if geography allows
      - Build a repository of examples and other artifacts that is easy to locate.

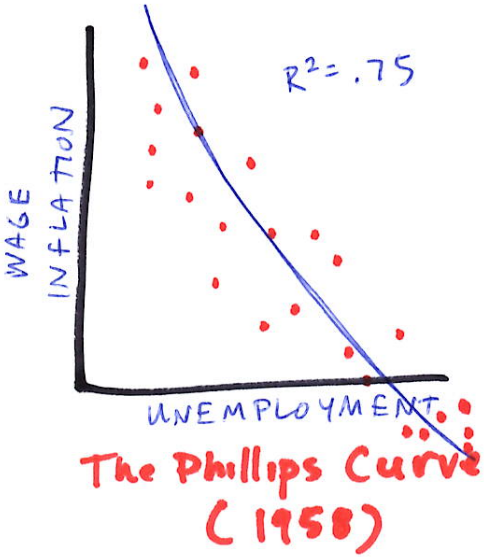
Conclusion

- Graduates of this program at U of A are some of the more likely candidates to become experts in one or more domains in their future companies!
  - This Applied Math program teaches students how to learn new disciplines quickly as well as how to apply fundamental math skills (anywhere).
- One of the best ways to demonstrate to your org that you are qualified for senior technical positions is to demonstrate the ability to build your own training programs.
  - One thing you'll learn (if you already haven't) is that you cannot be promoted if you cannot be replaced
  - And if you are working in a critical field that is growing, you will need to develop your next work force effectively

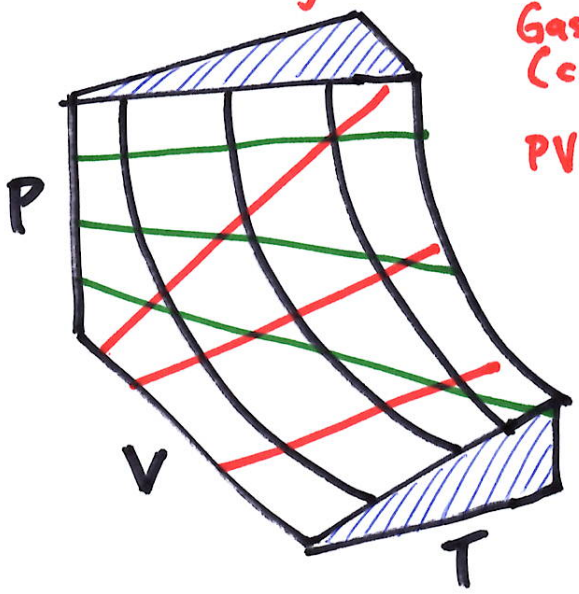
The Basic Math of ML

Def'n: Machine Learning is the science of creating machines/programs that improve from experience and interaction.

OLD IDEA: Curve Fitting



The Phillips Curve (1958)



Gas Law (c. 1800)

$PV = nRT$

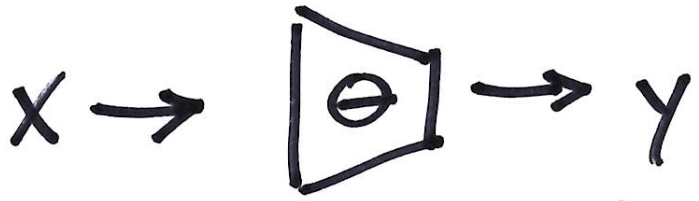
ML extends this technique in a computationally tractable way into high dimensions

Two things to note...

Loss Function - used to optimize the fit of the curve (model). Use Gaussian Least Square method (1795) along with numerous others

Metrics - help evaluate the fit after optimization. R^2 (Coeff. of Determination) describes the proportion of variance in y that is predictable from x

Finding a suitable function, given examples of Input/Output behavior



$\theta \in R^n$

Θ can be thought of as the set of parameters that generate (x, y)

So our loss function can be (simply) described as...

$$\ell(\theta) = \min \sum (\text{score} - \sum \theta)^2$$

(Desired output: score; Model's output: $\sum \theta$)

Note: Training = optimizing loss function

Testing = using trained model to predict y for new X

GRADIENT DESCENT = TECHNIQUE TO OPTIMIZE OUR LOSS FN

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla(\ell), \quad \eta \in \mathbb{R}$$

- $\theta^{(t+1)}$: NEXT POS'N
- $\theta^{(t)}$: CURRENT POS'N
- $-$: OPPOSITE DIRECTION
- η: STEP SIZE
- $\nabla(\ell)$: DIRECTION OF FASTEST INCREASE; LOSS

Convex Loss Functions

If we look carefully, our loss fn can be modeled as $y=x^2$

Use this process to find values of w and b that minimize $y = wx + b$

Math to do this

1) Simplify loss fn to

$$J_{w,b} = \frac{1}{N}\sum_{i=1}^{N}(\text{Error}_i)^2, \quad \text{Error} = Y' - Y$$

2) Simplify again by removing $\frac{1}{N}$ and summation (this works, but the Σ is important for batch G.D.)

$$J_{w,b} = (\text{Error}_i)^2 \therefore \frac{\partial J}{\partial w} = 2 \cdot \text{Error} \cdot \frac{\partial}{\partial w}\text{Error}$$

$$\frac{\partial J}{\partial b} = 2 \cdot \text{Error} \cdot \frac{\partial}{\partial b}\text{Error}$$

3) Take Partial Derivatives

$$\frac{\partial J}{\partial w} = 2 \cdot \text{Error} \cdot \frac{\partial}{\partial w}(mX + b - y) = 2 \cdot \text{Error} \cdot X$$

$$\frac{\partial J}{\partial b} = 2 \cdot \text{Error} \cdot \frac{\partial}{\partial b}(mX + b - y) = 2 \cdot \text{Error} \cdot 1$$

Back to $\theta^{t+1} = \theta^t - \eta \nabla(\ell)$, $\theta = w, b$

$$\frac{\partial J}{\partial w} = 2 \cdot \text{Error} \cdot X \qquad \frac{\partial J}{\partial b} = 2 \cdot \text{Error}$$

First, ignore the constant it will get absorbed in the learning rate when we do the next step...

now we know that $\eta \nabla(\ell) = \eta \frac{\partial J}{\partial w}$ for w

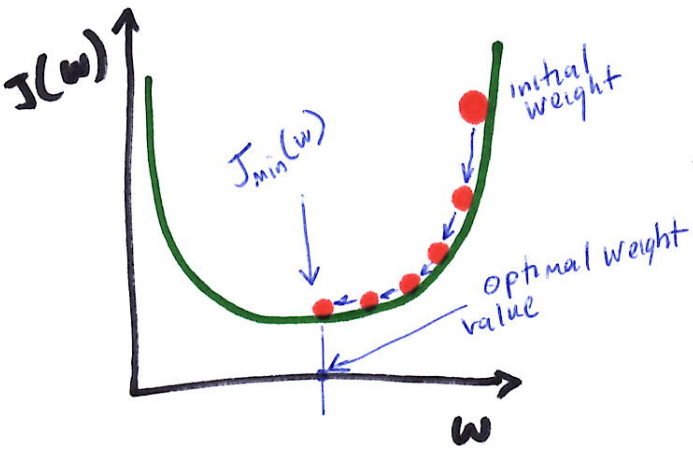
and $\eta \frac{\partial J}{\partial b}$ for b

Therefore for our weights, w:

$$w^{t+1} = w^t - \eta \cdot \text{Error} \cdot X$$

and for our biases:

$$b^{t+1} = b^t - \eta \cdot \text{Error}$$



RECAP

1) Take tangent from starting point
2) Find slope
3) ID how much change is required ($\frac{\partial J}{\partial w}$, $\frac{\partial J}{\partial b}$)
4) Multiply change ($\nabla(\ell)$ by learning rate
5) Calculate new θ parameter (w, b)

INTERESTING THOUGHT

Q: Why use Gradient Descent at all? Can't a closed form solution be used?

A: Sometimes the Closed form Analysis might be best, but in problems with high dimensionality ($\theta \in \mathbb{R}^n$), a Fourier Analysis, for example, needs computation of $O(e^n)$ and becomes computationally intractable. Gradient Descent's iterative approach is likely to be preferred (and it can also be easily parallelized)