

# Learning How to Deep Learn in a Not So Deep Way

*HackHW 2/17/18*

Mihail Eric

# Introduction

# Introduction

- Who am I?



Hackathon 5:07am

The weak have fallen — with [Mihail Eric](#)

# Introduction

- Who am I?

# Introduction

- Who am I?
  - Stanford BS in CS 2016, MS in CS 2017

# Introduction

- Who am I?
  - Stanford BS in CS 2016, MS in CS 2017
  - Concentration in Artificial Intelligence

# Introduction

- Who am I?
  - Stanford BS in CS 2016, MS in CS 2017
  - Concentration in Artificial Intelligence
    - Research in natural language understanding and dialogue systems with emphasis on deep learning techniques



# Introduction

- Who am I?
  - Stanford BS in CS 2016, MS in CS 2017
  - Concentration in Artificial Intelligence
    - Research in natural language understanding and dialogue systems with emphasis on deep learning techniques
  - Currently engineer at RideOS in San Francisco

# What is Deep Learning?

# What is Deep Learning?



deep learning



All

News

Videos

Books

**Images**

More

Settings

Tools

#84517548

# What is Deep Learning?

# What is Deep Learning?

- Artificial intelligence?

# What is Deep Learning?

- Artificial intelligence?
- Machine learning?

# What is Deep Learning?

- Artificial intelligence?
- Machine learning?
- umm... Alpha Go?



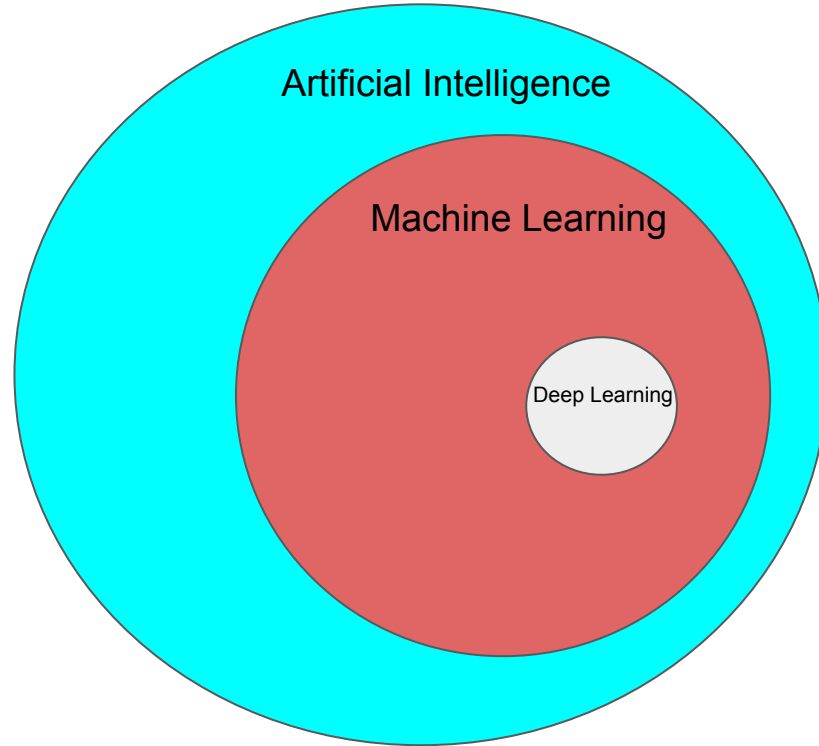
# What is Deep Learning?

- Artificial intelligence?
- Machine learning?
- umm... Alpha Go?
- Bitcoin?

# What is Deep Learning?

- Artificial intelligence?
- Machine learning?
- umm... Alpha Go?
- ~~Bitecoin?~~

# What is Deep Learning?

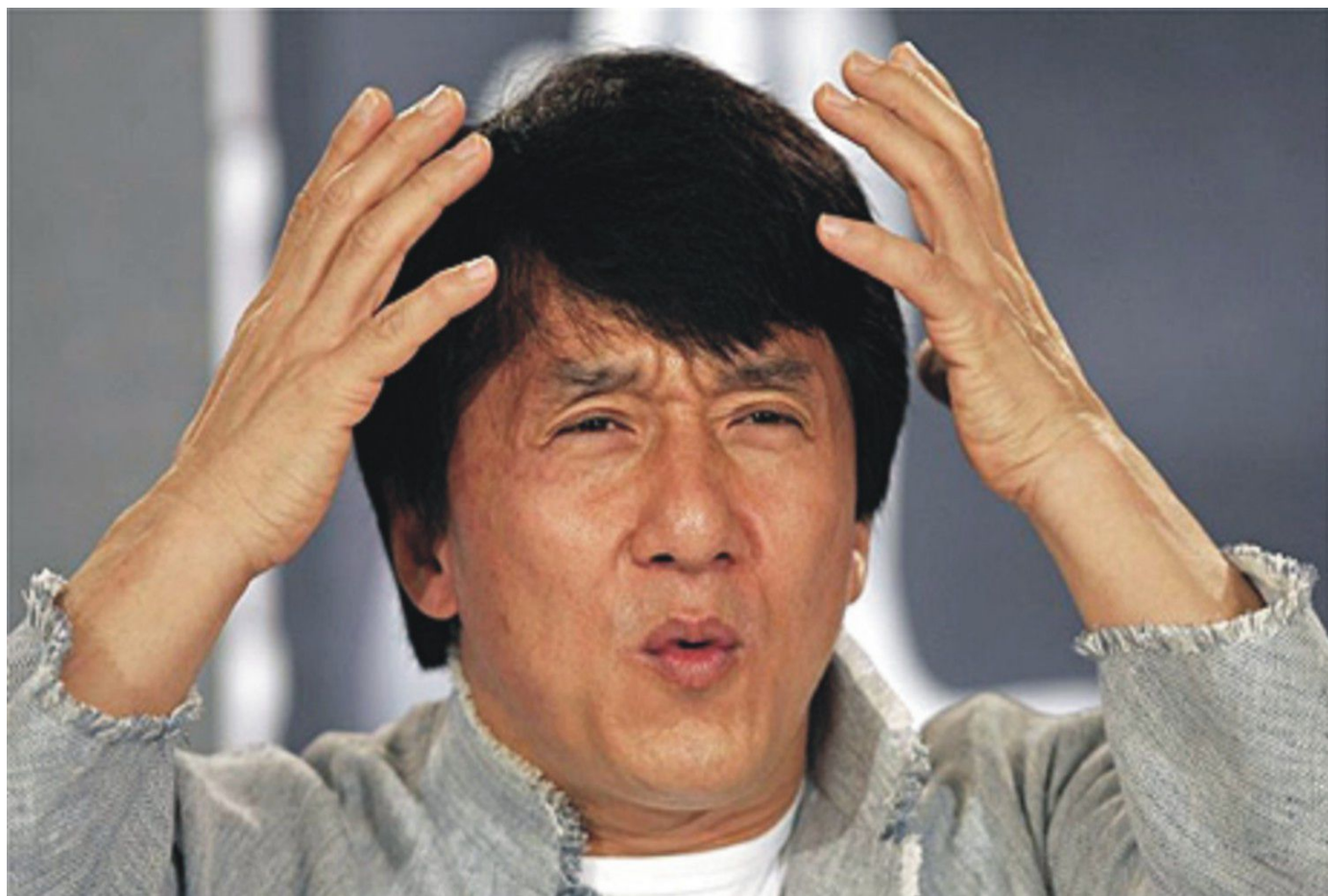


# Artificial Intelligence

- The science of developing techniques to endow non-sentient entities (machines) with complex reasoning abilities




















# Machine Learning

- The development of algorithms for learning statistical relationships in data that can be used in predictive tasks



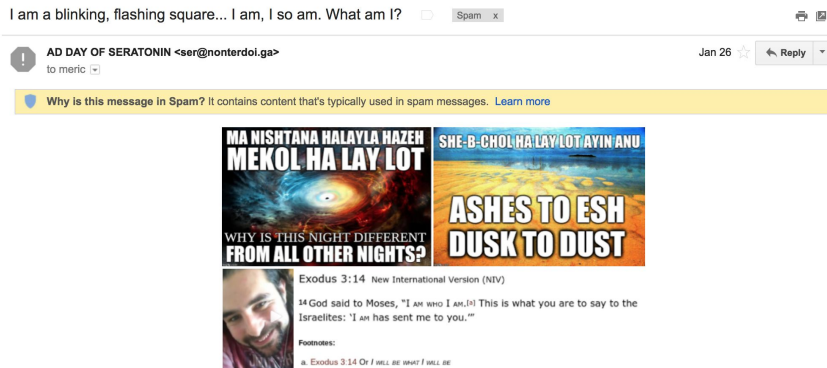
# Machine Learning

- Let's say you're trying to build a spam filter for some company... called Hoogle

From	Subject
 Adelaide Fatimah	a \$12000 watch, we sell at \$200, Quality watches at ...
 antonino rodney	Goodiest c1alis
 Irina Gidget	FDA Approved Medications: \$1.12/pill for Viagr...
 tom@messagingtime...	tom@messagingtimes.com, Up to 20% OFF
 Samantha Hickey	Enlarge, Widen and Strengthen
 churchill ravi	MSG #:19846 The world's largest online presc...
 abel yanjun	MSG #:84037 World's lowest prices on largest...
 Maureen Orr	Recapture a bit of your youth again
 nanako258@yahoo.c...	40□Î~È□ã,Å□S,à□gÎ,à-ü,â,³,ê,½,¢•û,Í[-ü,â,...
 Jerald Shook	a xmas gift to your wife is your bigger PE gs ft...
 Blanca Petty	Mit und schaffen Sie das was Frauen wollern!
 Lynne Mcneal	xp oem software
 emerson forrest	from Stella Vargas
 Revolution Jobs	Hundreds of digital careers on Revolution Jobs
 Auto Loan Department	GET APPROVED!
 jacquelyn	hi from jacquelyn
 ParkRoyalCancun	Visit Cancun With A 3 Night Free Stay - No Pur...
 Colon Cleanse Samples	View this LifeChanging Breakthrough
 o05689ok97@tom.com	40□Î~È□ã,Å□S,à□gÎ,à-ü,â,³,ê,½,¢•û,Í[-ü,â,...

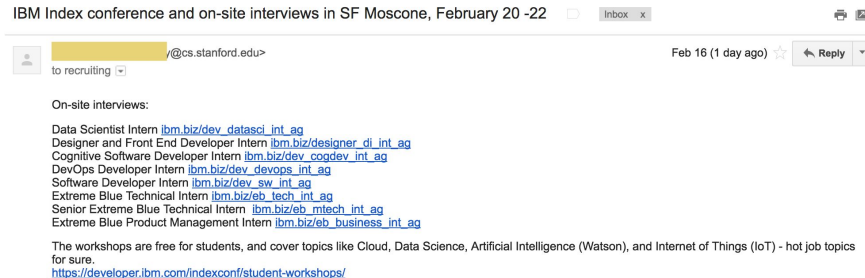


- You are given 1000 emails with labels per email indicating **SPAM** or **NOSPAM**



**cursor.**

**SPAM**



**NOSPAM**

# Identifying Spam

I am a blinking, flashing square... I am, I so am. What am I?

Spam x



AD DAY OF SERATONIN <ser@nonterdoi.ga>

to meric

Jan 26



Reply



Why is this message in Spam? It contains content that's typically used in spam messages. [Learn more](#)



**cursor.**

# Identifying Spam

I am a blinking, flashing square... I am, I so am. What am I?

Spam x



AD DAY OF SERATONIN <ser@nonterdoi.ga>

to meric

Jan 26



Reply



Why is this message in Spam? It contains content that's typically used in spam messages. [Learn more](#)



**cursor.**

- Grammatically incorrect subject

# Identifying Spam

I am a blinking, flashing square... I am, I so am. What am I?

Spam x



AD DAY OF SERATONIN <ser@nonterdoi.ga>

to meric

Jan 26



Reply



Why is this message in Spam? It contains content that's typically used in spam messages. [Learn more](#)



**cursor.**

- Grammatically incorrect subject
- @ domain is unrecognized given user history

# Identifying Spam

I am a blinking, flashing square... I am, I so am. What am I?

Spam x



AD DAY OF SERATONIN <ser@nonterdoi.ga>

to meric

Jan 26



Reply



Why is this message in Spam? It contains content that's typically used in spam messages. [Learn more](#)



**cursor.**

- Grammatically incorrect subject
- @ domain is unrecognized given user history
- A lot of memes in language never seen before in my email

# Identifying Non-Spam

IBM Index conference and on-site interviews in SF Moscone, February 20 -22

Inbox x



/@cs.stanford.edu>

Feb 16 (1 day ago) ☆

Reply ▾

to recruiting ▾

On-site interviews:

Data Scientist Intern [ibm.biz/dev\\_datasci\\_int\\_ag](#)  
Designer and Front End Developer Intern [ibm.biz/designer\\_di\\_int\\_ag](#)  
Cognitive Software Developer Intern [ibm.biz/dev\\_cogdev\\_int\\_ag](#)  
DevOps Developer Intern [ibm.biz/dev\\_devops\\_int\\_ag](#)  
Software Developer Intern [ibm.biz/dev\\_sw\\_int\\_ag](#)  
Extreme Blue Technical Intern [ibm.biz/eb\\_tech\\_int\\_ag](#)  
Senior Extreme Blue Technical Intern [ibm.biz/eb\\_mtech\\_int\\_ag](#)  
Extreme Blue Product Management Intern [ibm.biz/eb\\_business\\_int\\_ag](#)

The workshops are free for students, and cover topics like Cloud, Data Science, Artificial Intelligence (Watson), and Internet of Things (IoT) - hot job topics for sure.

<https://developer.ibm.com/indexconf/student-workshops/>



# Identifying Non-Spam

IBM Index conference and on-site interviews in SF Moscone, February 20 -22

Inbox x



/@cs.stanford.edu>

Feb 16 (1 day ago)



Reply

to recruiting

On-site interviews:

Data Scientist Intern [ibm.biz/dev\\_datasci\\_int\\_ag](#)

Designer and Front End Developer Intern [ibm.biz/designer\\_di\\_int\\_ag](#)

Cognitive Software Developer Intern [ibm.biz/dev\\_cogdev\\_int\\_ag](#)

DevOps Developer Intern [ibm.biz/dev\\_devops\\_int\\_ag](#)

Software Developer Intern [ibm.biz/sw\\_int\\_ag](#)

Extreme Blue Technical Intern [ibm.biz/eb\\_tech\\_int\\_ag](#)

Senior Extreme Blue Technical Intern [ibm.biz/eb\\_mtech\\_int\\_ag](#)

Extreme Blue Product Management Intern [ibm.biz/eb\\_business\\_int\\_ag](#)

The workshops are free for students, and cover topics like Cloud, Data Science, Artificial Intelligence (Watson), and Internet of Things (IoT) - hot job topics for sure.

<https://developer.ibm.com/indexconf/student-workshops/>



- Mention of IBM

# Identifying Non-Spam

IBM Index conference and on-site interviews in SF Moscone, February 20 -22

Inbox x



/@cs.stanford.edu>

Feb 16 (1 day ago) ☆

Reply ▾

to recruiting ▾

On-site interviews:

Data Scientist Intern [ibm.biz/dev\\_datasci\\_int\\_ag](#)  
Designer and Front End Developer Intern [ibm.biz/designer\\_di\\_int\\_ag](#)  
Cognitive Software Developer Intern [ibm.biz/dev\\_cogdev\\_int\\_ag](#)  
DevOps Developer Intern [ibm.biz/dev\\_devops\\_int\\_ag](#)  
Software Developer Intern [ibm.biz/dev\\_sw\\_int\\_ag](#)  
Extreme Blue Technical Intern [ibm.biz/eb\\_tech\\_int\\_ag](#)  
Senior Extreme Blue Technical Intern [ibm.biz/eb\\_mtech\\_int\\_ag](#)  
Extreme Blue Product Management Intern [ibm.biz/eb\\_business\\_int\\_ag](#)

The workshops are free for students, and cover topics like Cloud, Data Science, Artificial Intelligence (Watson), and Internet of Things (IoT) - hot job topics for sure.

<https://developer.ibm.com/indexconf/student-workshops/>



- Mention of IBM
- @cs.stanford.edu probably whitelisted



# Identifying Non-Spam

IBM Index conference and on-site interviews in SF Moscone, February 20 -22

Inbox x



/@cs.stanford.edu>

Feb 16 (1 day ago)



Reply

to recruiting

On-site interviews:

Data Scientist Intern [ibm.biz/dev\\_datasci\\_int\\_ag](#)  
Designer and Front End Developer Intern [ibm.biz/designer\\_di\\_int\\_ag](#)  
Cognitive Software Developer Intern [ibm.biz/dev\\_cogdev\\_int\\_ag](#)  
DevOps Developer Intern [ibm.biz/dev\\_devops\\_int\\_ag](#)  
Software Developer Intern [ibm.biz/dev\\_sw\\_int\\_ag](#)  
Extreme Blue Technical Intern [ibm.biz/eb\\_tech\\_int\\_ag](#)  
Senior Extreme Blue Technical Intern [ibm.biz/eb\\_mtech\\_int\\_ag](#)  
Extreme Blue Product Management Intern [ibm.biz/eb\\_business\\_int\\_ag](#)

The workshops are free for students, and cover topics like Cloud, Data Science, Artificial Intelligence (Watson), and Internet of Things (IoT) - hot job topics for sure.

<https://developer.ibm.com/indexconf/student-workshops/>



- Mention of IBM
- @cs.stanford.edu probably whitelisted
- Mention of on-site interviews, students, Watson, artificial intelligence, etc.

# Spam Filtering

# Spam Filtering

- You combine your insights to identify and extract a set of *features* from the data

# Spam Filtering

- You combine your insights to identify and extract a set of *features* from the data
- When you see an email mention IBM that isn't spam, you +1 to the CONTAINS\_IBM feature

# Spam Filtering

- You combine your insights to identify and extract a set of *features* from the data
- When you see an email mention IBM that isn't spam, you +1 to the CONTAINS\_IBM feature
- When you see a spam email contain memes in a foreign language, you +1 to the FOREIGN\_LANGUAGE feature

# Testing Your System's Performance

# Testing Your System's Performance

- Do you test on the 1000 emails you used to engineer features?

# Testing Your System's Performance

- Do you test on the 1000 emails you used to engineer features?
  - NO!!!



# Testing Your System's Performance

- Do you test on the 1000 emails you used to engineer features?
  - NO!!!
  - Why?

# Testing Your System's Performance

- Do you test on the 1000 emails you used to engineer features?
  - NO!!!
  - Why?
- We need a set of emails we **haven't** seen before

# Testing Your System's Performance

- Do you test on the 1000 emails you used to engineer features?
  - NO!!!
  - Why?
- We need a set of emails we **haven't** seen before
  - Test on 100 new emails

# Testing Your System's Performance

- Do you test on the 1000 emails you used to engineer features?
  - NO!!!
  - Why?
- We need a set of emails we **haven't** seen before
  - Test on 100 new emails
  - Held-out test set

# Summary

# Summary

- Paradigm of training from labelled data, testing on unseen data underlies almost all of machine learning!

# Summary

- Paradigm of training from labelled data, testing on unseen data underlies almost all of machine learning!
  - Predicting housing prices

# Summary

- Paradigm of training from labelled data, testing on unseen data underlies almost all of machine learning!
  - Predicting housing prices
  - Classifying images



# Summary

- Paradigm of training from labelled data, testing on unseen data underlies almost all of machine learning!
  - Predicting housing prices
  - Classifying images
  - Building a speech recognizer

# Summary

- Paradigm of training from labelled data, testing on unseen data underlies almost all of machine learning!
  - Predicting housing prices
  - Classifying images
  - Building a speech recognizer
  - Predicting whether someone should be a friend recommendation in FB

# Summary

- Paradigm of training from labelled data, testing on unseen data underlies almost all of machine learning!
  - Predicting housing prices
  - Classifying images
  - Building a speech recognizer
  - Predicting whether someone should be a friend recommendation in FB
  - etc.

Back to Deep Learning...

# Back to Deep Learning...

- Ideas have been around since the 1980s

# Back to Deep Learning...

- Ideas have been around since the 1980s
- Resurgence of interest in 2011

# Back to Deep Learning...

- Ideas have been around since the 1980s
- Resurgence of interest in 2011
  - A deep neural network achieved ~11% error reduction on ImageNet dataset

# Back to Deep Learning...

- Ideas have been around since the 1980s
- Resurgence of interest in 2011
  - A deep neural network achieved ~11% error reduction on ImageNet dataset
- Why now?



# Back to Deep Learning...

- Ideas have been around since the 1980s
- Resurgence of interest in 2011
  - A deep neural network achieved ~11% error reduction on ImageNet dataset
- Why now?
  - Better techniques for training neural networks

# Back to Deep Learning...

- Ideas have been around since the 1980s
- Resurgence of interest in 2011
  - A deep neural network achieved ~11% error reduction on ImageNet dataset
- Why now?
  - Better techniques for training neural networks
  - More data

# Back to Deep Learning...

- Ideas have been around since the 1980s
- Resurgence of interest in 2011
  - A deep neural network achieved ~11% error reduction on ImageNet dataset
- Why now?
  - Better techniques for training neural networks
  - More data
  - Better hardware

Why is it different?

# Why is it different?

- Previously when tackling spam filtering, **feature engineering** was the crux of building the best system

# Why is it different?

- Previously when tackling spam filtering, **feature engineering** was the crux of building the best system
  - Manual domain-specific characteristics of data

# Why is it different?

- Previously when tackling spam filtering, **feature engineering** was the crux of building the best system
  - Manual domain-specific characteristics of data
- Imagine if all we had was 1000 emails with labels and the model learned **EVERYTHING** on its own

IBM Index conference and on-site interviews in SF Moscone, February 20 -22

Inbox x



/@cs.stanford.edu>

to recruiting

Feb 16 (1 day ago)

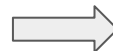
Reply

On-site interviews:

Data Scientist Intern [ibm.biz/dev\\_datasci\\_int\\_ag](#)  
Designer and Front End Developer Intern [ibm.biz/designer\\_di\\_int\\_ag](#)  
Cognitive Software Developer Intern [ibm.biz/dev\\_cogdev\\_int\\_ag](#)  
DevOps Developer Intern [ibm.biz/dev\\_devops\\_int\\_ag](#)  
Software Developer Intern [ibm.biz/dev\\_sw\\_int\\_ag](#)  
Extreme Blue Technical Intern [ibm.biz/eb\\_tech\\_int\\_ag](#)  
Senior Extreme Blue Technical Intern [ibm.biz/eb\\_mtech\\_int\\_ag](#)  
Extreme Blue Product Management Intern [ibm.biz/eb\\_business\\_int\\_ag](#)

The workshops are free for students, and cover topics like Cloud, Data Science, Artificial Intelligence (Watson), and Internet of Things (IoT) - hot job topics for sure.

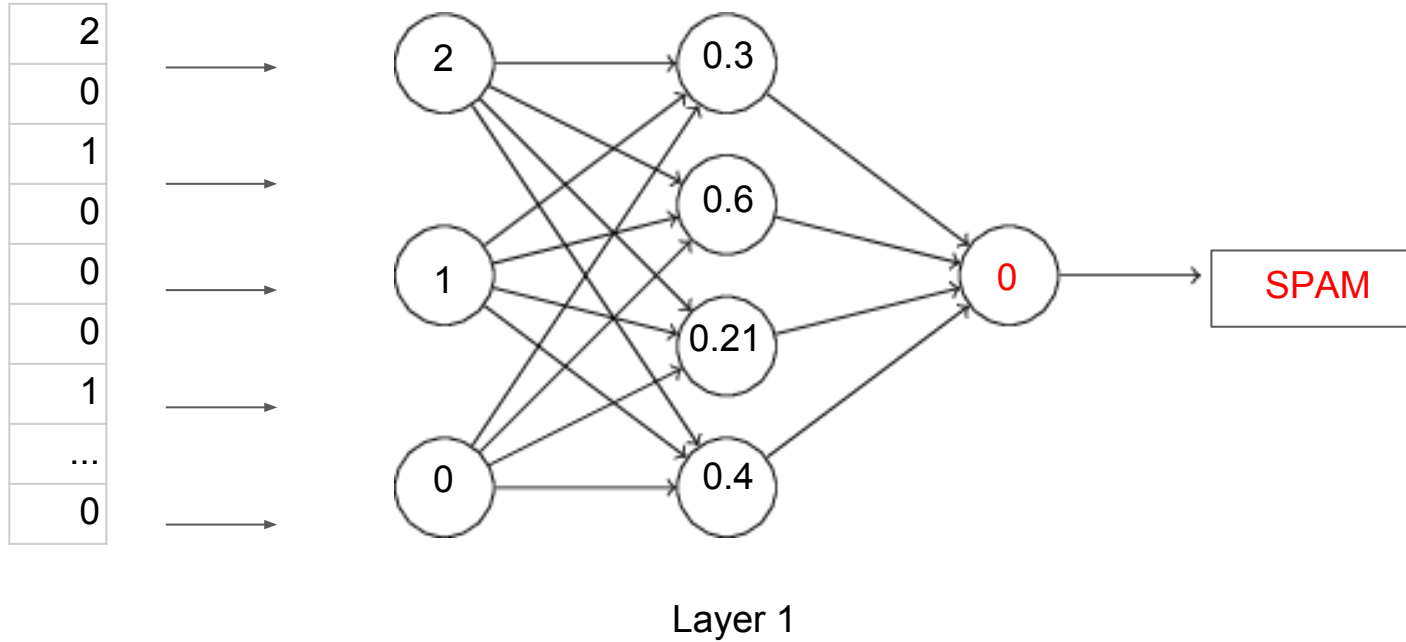
<https://developer.ibm.com/indexconf/student-workshops/>



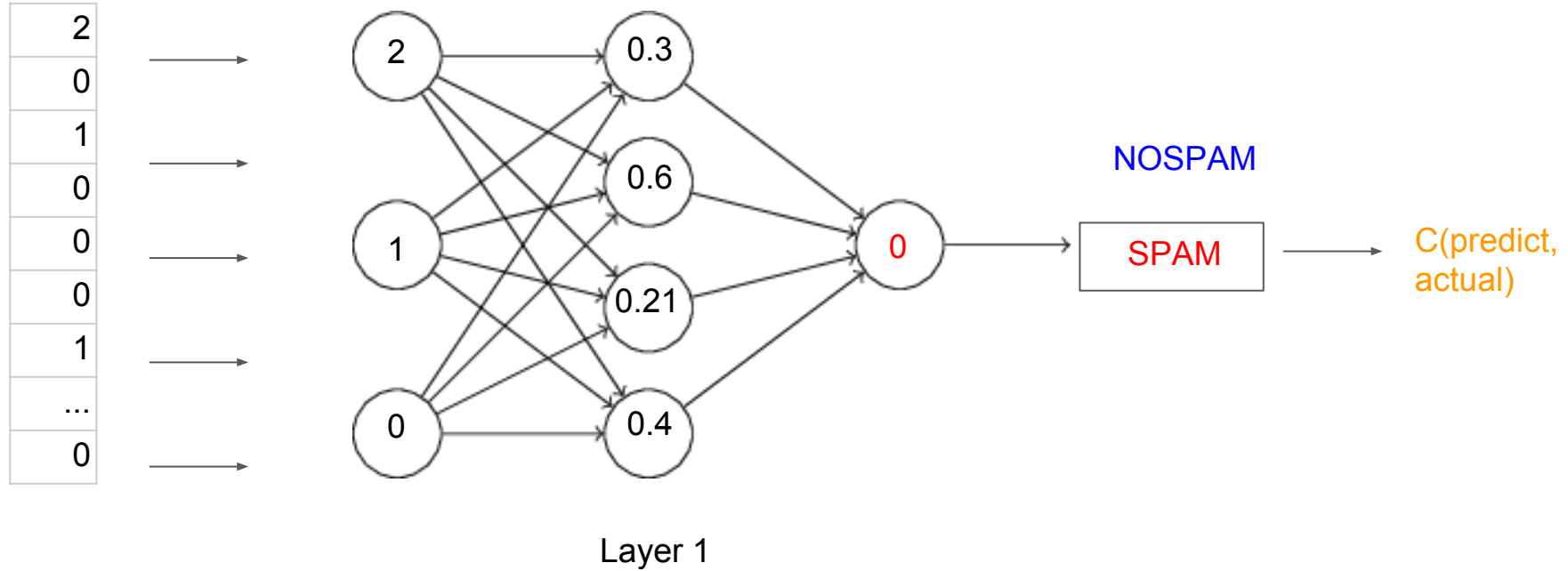
2
0
1
0
0
0
1
...
0



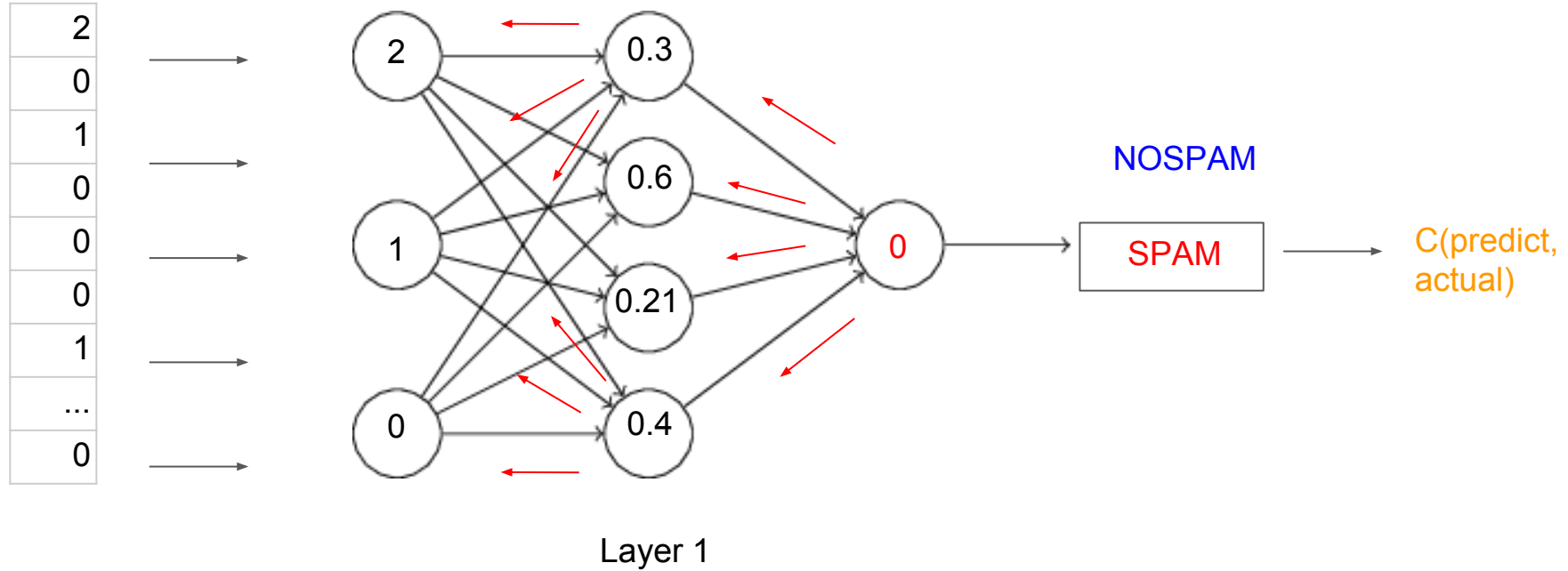
# Training the Network



# Training the Network



# Training the Network



Lather, Rinse, Repeat!

# Let's train a neural network!

<https://github.com/mihail911/deep-hwack>

# Thanks For Listening!

[mihail911@gmail.com](mailto:mihail911@gmail.com)

[www.mihaileric.com](http://www.mihaileric.com)