

“Machine Learning” and “A Personal Journey for Engineering”

Yanjun (Jane) Qi, PhD
Department of Computer Science,
University of Virginia

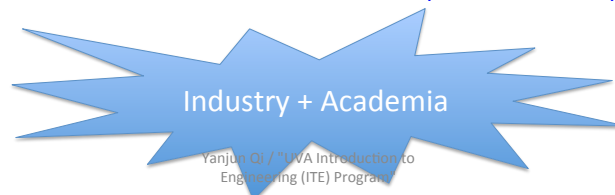
2015.06.23

6/24/15

Yanjun Qi / "UVA Introduction to Engineering (ITE) Program"

About Me

- **Education:**
 - PhD from School of Computer Science, [Carnegie Mellon University](#) (@ Pittsburgh, PA)
 - BS in Department of Computer Science, [Tsinghua Univ.](#) (@ Beijing, China)
- **Research interests:**
 - **Machine Learning, Bioinformatics, Data Mining, Biomedical Informatics**
- **Five Years' of Industry Research Lab in the past :**
 - 2008 summer – 2013 summer, [Researcher in Machine Learning](#) Department, NEC Labs America (@ Princeton, NJ)
 - 2013 Fall – Present, Assistant Professor, [Computer Science Dept., UVA](#)



2
6/24/15

Yanjun Qi / "UVA Introduction to Engineering (ITE) Program"

ROADMAP

- Machine Learning & Applications
- My Experience: A Personal Journey of Engineering Study

6/24/15

YanJun Qi / "UVA Introduction to
Engineering (ITE) Program"

OUR DATA-RICH WORLD



- Biomedicine
 - Patient records, brain imaging, MRI & CT scans, ...
 - Genomic sequences, bio-structure, drug effect info, ...
- Science
 - Historical documents, scanned books, databases from astronomy, environmental data, climate records, ...
- Social media
 - Social interactions data, twitter, facebook records, online reviews, ...
- Business
 - Stock market transactions, corporate sales, airline traffic, ...
- Entertainment
 - Internet images, Hollywood movies, music audio files, ...

6/24/15

YanJun Qi / "UVA Introduction to
Engineering (ITE) Program"

BIG DATA CHALLENGES

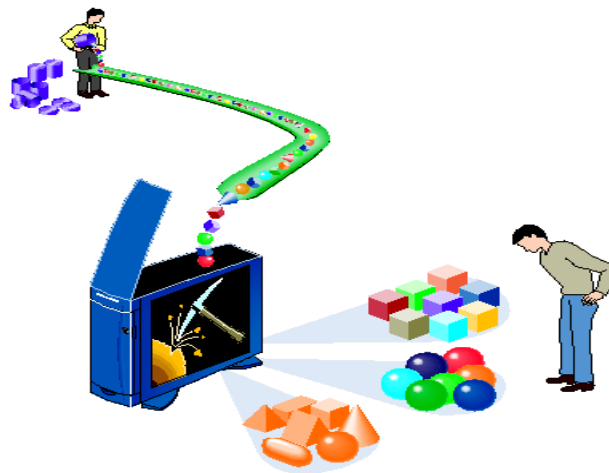
- Data capturing (sensor, smart devices, medical instruments, et al.)
- Data transmission
- Data storage
- Data management
- High performance data processing
- Data visualization
- Data security & privacy (e.g. multiple individuals)
-

- Data analytics
 - How can we analyze this big data wealth ?
 - E.g. Machine learning

6/24/15

Yanjun Qi / "UVA Introduction to
Engineering (ITE) Program"

Drowning in data, Starving for knowledge



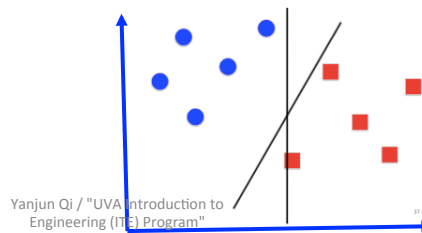
6/24/15

Yanjun Qi / "UVA Introduction to
Engineering (ITE) Program"

6

BASICS OF MACHINE LEARNING

- “The goal of machine learning is to build computer systems that can **learn and adapt from their experience.**” – Tom Dietterich
- “**Experience**” in the form of available **data examples** (also called as instances, samples)



6/24/15

Yanjun Qi / "UVA introduction to Engineering (ITE) Program"

e.g. SUPERVISED LEARNING

- Find function to map **input** space X to **output** space Y $f : X \rightarrow Y$
- So that the **difference** between y and $f(x)$ of each example x is small.

e.g.

x	I believe that this book is not at all helpful since it does not explain thoroughly the material . it just provides the reader with tables and calculations that sometimes are not easily understood ...
----------	--

y	-1
----------	----

Output Y: {1 / Yes , -1 / No }
e.g. Is this a positive product review ?

Input X : e.g. a piece of English text

6/24/15

Yanjun Qi / UVA CS 4501-01-6501-07

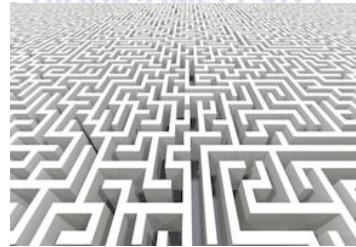
8

BIG DATA CHALLENGES FOR MACHINE LEARNING

LARGE-SCALE



HIGH-COMPLEXITY



6/24/15

Yanjun Qi / "UVA Introduction to Engineering (ITE) Program"

Large-Scale Machine Learning: SIZE MATTERS

LARGE-SCALE



- One thousand data instances
- One million data instances
- One billion data instances
- One trillion data instances

Those are not different numbers,
those **are different mindsets !!!**

6/24/15

Yanjun Qi / "UVA Introduction to Engineering (ITE) Program"

RELATED DISCIPLINES

- **Artificial Intelligence**
- Data Mining
- Probability and Statistics
- Information theory
- Numerical optimization
- Computational complexity theory
- Control theory (adaptive)
- Psychology (developmental, cognitive)
- Neurobiology
- Linguistics
- Philosophy

6/24/15

Yanjun Qi / "UVA Introduction to
Engineering (ITE) Program"

How can we build more intelligent
computer / machine ?



R2-D2 and C-3PO

@ Star Wars – 1977

to serve human beings,
and
fluent in "over six million
forms of communication"

12

How can we build more intelligent computer / machine ?



Jeopardy Game
 → Requires a Broad Knowledge Base
 6/24/15

IBM Watson

→ an artificial intelligence computer system capable of answering questions posed in natural language developed in IBM's DeepQA project.



IBM WATSON

How can we build more intelligent computer / machine ?

Apple Siri → an intelligent **personal assistant** and knowledge navigator



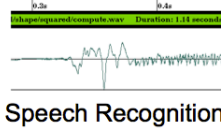
How may I help you,
human?

Machine Learning is changing the WORLD

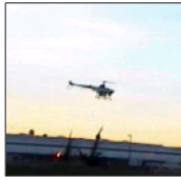
Data:

PatientID	Age	Parity	Emergency C-Section
1	28	0	Yes
2	32	1	No
3	25	0	Yes
4	35	2	No
5	29	1	Yes
6	31	0	No
7	27	1	Yes
8	33	0	No
9	26	1	Yes
10	34	0	No

Mining Databases



Speech Recognition



Control learning



Object recognition

Text analysis

Peter H. van Oopjen has served as **Director of the Center for Quality Improvement** since its acquisition by Interpoint in 1994 and a Director of ADIG since 1996. Until its acquisition by Crane Co. in October 1996, **Mr. van Oopjen** served as **Director of Quality** at **Price Waterhouse LLP** and at Bain & Company in Boston and London. He has additional experience in medical electronics and venture capital. **Mr. van Oopjen** also serves as a **Director** of **Spacelabs Medical, Inc.** He holds a B.A. from Whitman College and an M.B.A. from Harvard Business School, where he was a **Baker Scholar**.

Many more !

YanJun Qi / "UVA Introduction to Engineering (ITE) Program"

ROADMAP

- Machine Learning & Applications
- My Experience: A Personal Journey of Engineering Study



To Share Some Lessons learned / Choices made

➔ my personal journey of studying engineering



A bit more about me

- Hometown:
 - northwest China (less developed region)
- Family:
 - Working class family (loving & caring)
 - My parents know nothing about college / graduate study / high-tech

I. Before College:

- Suggestion: Do not listen to certain comments, e.g., Girls are not good at math / Engineering / Science !



IGNORE WRONG
COMMENTS !!!

- Focusing on academic study

6/24/15

Yanjun Qi / "UVA Introduction to
Engineering (ITE) Program"

19

2. During College

- Just like the Olympics: “the thrill of victory and the agony of defeat!”
- I have learned that → failure is a good thing for your growth !



Improve Yourself !
(Professionally &
Personally)

6/24/15

Yanjun Qi / "UVA Introduction to
Engineering (ITE) Program"

20

3. During Graduate Study

- Suggestion: Push the limit of yourself as much as you can
- Time management is KEY !



Improve Yourself !
(professionally)


6/24/15

YanJun Qi / "UVA Introduction to
Engineering (ITE) Program"

21

4. After PhD:

- Love to continue research
- Industry Research vs. Academic Research ?
- → So I did both !!!



Aim to Use
Technology to change
the world

6/24/15 YanJun Qi / "UVA Introduction to
Engineering (ITE) Program"

22

5. During PhD and after

- Work / Life Balance
- Working mommy moments



Summary:

Tailor Goals w.r.t. Experiences

- Why are you here ?
- What do you want out of this program / job / education ?
- How can you get it ?

Take-Home Message / Question?

- Question to Ask Yourself ?
 - What makes you feel very excited ?
 - What makes you feel “wow... my work has impact !”
 - What makes you feel “I am so proud of myself !”

- How do you picture yourself in 5 years, 10 years, 20 years, ... ??

Yanjun Qi / "UVA Introduction to Engineering (ITE) Program"

25

Feel the fear and do it anyway!



“Confidence does not require perfection”

Great advice from CRA-W:

6/24/15

Yanjun Qi / "UVA Introduction to Engineering (ITE) Program"

26

Questions

