

Deep Learning in Computer Vision and NLP

Liangliang Cao
James Fan

<http://llcao.net/cu-deeplearning15/>



Introduction



James J. Fan

RSM - DeepQA (Watson) Project
Thomas J. Watson Research Center, Yorktown Heights, NY USA
fanj@us.ibm.com +1-914-945-2493

Profile Publications Patents

Project Pages

• [The DeepQA Research Team](#)

Professional Interests

- [Computer Science](#)
- [Artificial Intelligence](#)
- [Natural Language Processing](#)
- [Web](#)

James Fan is a research staff member at IBM Research. His research interests include question answering, knowledge representation and reasoning, natural language processing and machine learning. James is currently working [the DeepQA](#) project which is advancing the state-of-the-art in automatic, open domain question answering technology. The DeepQA team is pushing question answering technology to levels of performance previously unseen and demonstrate the technology by playing Jeopardy! at the level of a human champion.



Liangliang Cao

Research Staff Member- Large Scale Computer Vision and New Media
Rm16-220, 1101 Kitchawan Rd, Yorktown Heights, NY 10598
liangliang.cao@us.ibm.com +1-914-945-2145

Profile Research Awards Professional Service

Project Pages

• [IBM Multimedia Analysis and Retrieval System \(IMARS\)](#)

Professional Interests

- [Knowledge Discovery and Data Mining](#)
- [Multimedia](#)
- [Natural Language Processing](#)

Research interests:

I am interested in research problems related to computer vision and social media. More specifically, I am curious with the the following questions:

- Can computer vision outperform human vision?
- Can a computer have surprising discovery from social media which is unknown to human being?
- Can a computer answer questions as reliably as human experts?

Teaching Assistant (to be confirmed)

- Colin Abraham Raffel



<http://colinraffel.com/>

You have to reconstruct his face when you interact with him online (say, Google group)

Course Website and Google Group

- Course schedule and slides will be put on the website
- “Resource” page under construction.

<http://llcao.net/cu-deeplearning15/>

Columbia University EECS E6894, Spring 2015

Deep Learning for Computer Vision and Natural Language Processing

Schedule

Reading

Resource

Final Projects

Instructors

- [Liangliang Cao](#) (liangliang.cao_at_gmail_dot_com)
- [James Fan](#) (jfan.us_at_gmail_dot_com)

Course Introduction

This graduate level research class focuses on deep learning techniques for vision and natural language processing problems. It gives an overview of the various deep learning models and techniques, and surveys recent advances in the related fields. This course uses [Theano](#) as the main programming tool. GPU programming experience are preferred although not required. Frequent paper presentations and a heavy programming workload are expected.

Announcement

First class will start on Jan 21.

Course Time and Location

Wednesday evening, 7:00-9:30pm, Classroom to be announced

Office Hour vs. Google Group

- No fixed office hour
 - But you are welcome to stay after the course and ask questions

- Google group of this course

<https://groups.google.com/forum/#!forum/ee6894-deep-learning>

Course Content and Schedule

- Introduction to Watson Jeopardy! and Deep Learning
- Programming Guide
- Deep Learning and Computer Vision
- Deep Learning and NLP
- Student Presentations and Course Projects

Why We Choose This Schedule

- Why Theano instead of other deep learning toolkit?
 - Theano can be used with both GPU and CPU
 - Theano can be used for both vision and NLP (or other apps)
 - Limitations of Theano
- How to give you enough hands-on experience?
 - Quick programming tour
 - In class programming competition
 - Course projects (team size to be announced)
 - In class student presentations

Grading

- 60% project
- 30% paper presentation
- 10% participation

Detailed Schedule

<http://llcao.net/cu-deeplearning15/>

Course Schedule

Part I: Background and Introduction

Week	Topic	Note
1 (1/21)	Liangliang Course overview James From deep QA to deep NLP: the success of IBM Jeopardy! and beyond	First homework assigned
2 (1/28)	Liangliang A computational viewpoint for deep learning Discussion of student project ideas	First homework due

Part II: Programming Guidance

Week	Topic	Note
3 (2/4)	James Quick tour of Theano programming	In class programming competition
4 (2/11)	Liangliang Comparing MLP and CNN with dropout for handwriting digit recognition	In class programming competition
5 (2/18)	Student Projects Mid-term project proposal presentation	

Part III: Deep Learning and Vision

Week	Topic	Note
6 (2/25)	Liangliang ImageNet large scale visual recognition Video recognition, action detection	In class student presentation
7 (3/4)	Liangliang Face: recognition and verification Object detection and face detection	In class student presentation
8 (3/11)	Liangliang OCR LSTM and related techniques	In class student presentation
No class (3/18)	Spring break	

Projects

What kind of projects would you like to take in this class?

1. NLP
2. Vision
3. NLP + Vision
4. Your own data or problem?

In the next class we will listen to your ideas and suggest some project ideas.

Project Requirements

One group of 3 students should finish their project with the following checkpoints:

- 2/18: Midterm project proposal
 - Every team prepares a 10 min presentation
 - Slides due on 2/17
- 4/22 and 4/29: Final project presentation
 - Every team prepares a 20-30 min presentation
 - Slides due on 4/21
- 5/10: Final project report due
 - 4+ pages report, in the format of top conferences (e.g., ACL or ICCV)

Student Paper Presentation

- Student should prepare to present the papers in one of the following topics:
 1. Word embedding
 2. Syntax and semantics
 3. NLP applications
 4. Large Scale Image/Video
 5. Face
 6. OCR
- We encourage students choose a topic which is related to their project
- Register presentation slots between Jan 28 and Feb 18 (First come first fill)

Be A Good Collaborator and An Honest Learner

- We hope you enjoy this class as a team learning process
 - Be devoted in collaboration
 - Be on time for discussion or delivery
 - Learn from your teammates
- Please do not copy homework/slides/reports
 - It is easy to find out by your peers or search engines
 - You may have trouble with university rules
 - And you lose the chance of learning or doing by yourself

Course Requirement

- Knowledgeable about NLP and/or vision and/or machine learning
- Fluent in Python and Numpy programming

Students will be strongly advised to drop the class if they cannot finish the first homework on time.

First Homework

Submit by Emailing

- liangliang.cao@gmail.com
- jfan.us@gmail.com

before the 2nd class (Jan. 28)