

Introduction to Deep Learning

Course Logistics

Spring 2023

Swathi Jadav

Outline

- Who should take this course?
- Who should not take this course?
- Learning Outcomes and Content Overview
- Course Logistics
- Scores, Cutoffs, and Grades
- Student and TA Expectations
- Teamwork, Study Groups, and Cheating
- Challenges

Outline

- Who should take this course?
- Introduction
- Learning Outcomes and Content Overview
- Course Logistics
- Scores, Cutoffs, and Grades
- Student and TA Expectations
- Teamwork, Study Groups, and Cheating
- Challenges

Who is this course designed for?

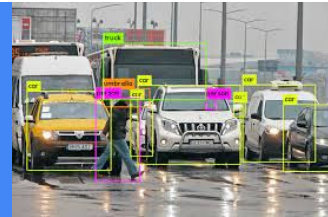
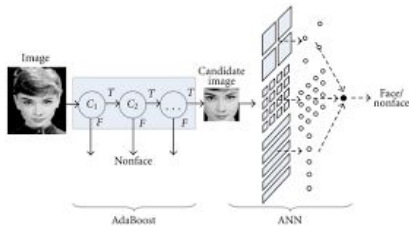
- Students from any background that want to learn deep learning
- Students who are willing to put in **20 hours** a week on this course
- Students who give continuous feedback and engage on Piazza
- Students who are mature and want to be challenged
- Students who want to be ready for AI research & engineering roles

Neural Networks are taking over!

- What are neural networks in simple words?
- A neural network is a method in artificial intelligence that teaches computers to **process data in a way that is inspired by the human brain**
- Neural Networks have become one of the major thrust areas recently in various pattern recognition, prediction, and analysis problems
- Established state-of-the-art in many problems, often exceeding the previous benchmarks by large margins

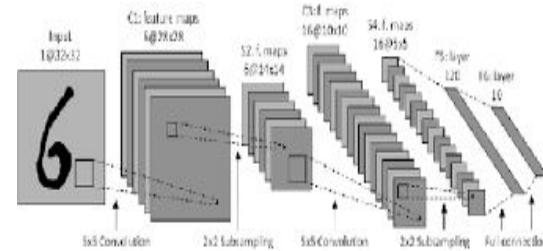
Breakthrough with Neural Networks

- Vision/ Perception - Face Detection, Face Recognition, Object Detection, Semantic Segmentation, Pose Estimation, Activity Recognition, Emotion Detection, Biometrics and forensics.



Breakthrough with Neural Networks

- Audio- Voice Assistants, ChatBots, Speech Recognition, machine Translation
- Text - Handwriting Recognition, text to speech recognition



Breakthrough with Neural Networks

- Robotics- Perception, Navigation, Guidance and control, Grasp Detection, Robot manipulation.



Breakthrough with Neural Networks

- Self-Driving Cars- Behaviour Prediction, Interpreting human actions, Task planning, Scene Understanding, Sensor Fusion, End-to-End Control



Most Recent

- DALL-E 2
- ChatGPT



Outline

- Who should take this course?
- Introduction
- Course Objectives
- Course Logistics
- Scores, Cutoffs, and Grades
- Student and TA Expectations
- Teamwork, Study Groups, and Cheating
- Challenges

Preparation For The Course

- Required Classes:
 - Fundamentals of Programming (or equivalent) - Python, Pytorch
- Highly Recommended Classes:
 - Calculus and Linear Algebra
- Nice To Have:
 - Vector Calculus and Software Engineering

Course Objectives

- Understanding neural networks
- Comprehending the models that do the previously mentioned tasks
 - And maybe build them
- Fearlessly design, build and train networks for various tasks

Course objectives

Concepts

- Some historical perspective
- Types of neural networks and underlying ideas
- Learning in neural networks
- Training, concepts, practical issues
- Architectures and applications

Practical

- Familiarity with training neural networks
- Implement various neural network architectures
- Implement state-of-art solutions for some problems

Overall

- Set you up for further research/work in your research area.

Course objectives

- Basic network formalisms:
 - Multilayer Perceptrons
 - Convolutional Neural Networks
 - Recurrent Neural Networks
 - Boltzmann Machines
- Some advanced formalisms
 - Generative models: VAEs
 - Adversarial models: GANs
 - Graph Neural Networks
- Topics we will touch upon:
 - Computer vision: recognizing images
 - Text processing: modelling and generating language
 - Machine translation: Sequence to sequence modelling
 - Modelling distributions and generating data
 - Speech recognition

Course Philosophy

- Our goal is to **teach** you deep learning. You are **graded** on your ability to **show** you understand deep learning.
- We give you many many opportunities to **demonstrate** your understanding. You will have multiple opportunities to **work** towards an A and **improve** on your shortcomings IF you are willing to put in the work.
- We tell you when you are wrong, why we think you are wrong, and **how you should think** about getting the right answer.
- We don't just want you to have deep learning knowledge, we want you to be able to **think critically and creatively** about designing solutions to new problems.

Reading Material

- List of books mentioned on the course webpage
- Additional reading material will also appear on the course pages

Instructors and TA's

Instructors:

- Bhiksha Raj
- Rita Singh

Teaching Assistants:

- TA information provided on the course website
- TA's present in the following Campuses:
 - ❖ Pittsburgh, Silicon Valley, kigali, Doha Campus

TA Office Hours:

- Remote and In-person TA Office Hours
- Office hour's schedule listed on the course page
- <https://deeplearning.cs.cmu.edu/S23/index.html>



Logistics

- *In-class lectures (unless rules change due to covid)*
 - Lectures are also recorded
 - Recordings will be posted
- Important that you view the lectures
 - Even if you think you know the topic
 - Your marks depend on viewing lectures
 - We will monitor attendance (more on this later)

Additional Logistics

- Reading and Engaging on Piazza is Mandatory You will fail quickly if you do not use Piazza.

Compute Infrastructure

- AWS Coupons would be provided
- Google Cloud Platform (GCP): 300\$ Worth free credits available
- Pittsburgh Supercomputing (PSC) access would be given

Attendance

- You get marks for attendance
 - Our performance metrics over the semesters show a distinct correlation between attendance and course scores
 - We also note a distinct *inverse* correlation between attendance and the amount of help you require on piazza and during office hours
 - To encourage attendance, we assign 1 mark for attendance
 - 1% of your total grade for 11685/11785/18786
 - 1.3333% of your total grade for 11485
 - This can be the difference between a B and an A
- We will track lecture attendance
 - More on next slide

Attendance

- You must either attend the lectures in person, or the streamed video
 - With some exceptions: see below
 - **We get attendance stats from your participation in in-class polls**
 - Polls will be conducted via Piazza. It is recommended to use the Piazza app available both on Android and iOS. You can also use Piazza in browser if you are attending via Zoom.
 - Use of electronic devices during in-person lectures is not permitted except during polls.
- Students in SV, and those stuck in bad timezones (i.e. if your local time is before 8am or after 5pm for the class) may alternately watch recorded lectures on mediatech instead
 - Mediatech records who watched and for how long. You must watch at least 70 mins of the lecture
 - If viewed on mediatech, the lectures of each week must be viewed before 8AM of the Monday following the following week
 - Otherwise, it doesn't count
- At the end of the semester, we will select a random subset of 50% of the lectures and tabulate attendance
- If you have attended at least 70% of these (randomly chosen) lectures, you get the attendance point **(Subject to change)**

Lecture Schedule

The lecture schedule is mentioned on the website.

The Slides and other materials are posted on the website.

Guest Lectures :

- To be Decided

Recitations

- 14 recitations (Excluding HW Bootcamps)
- Every Friday of the semester
- Will cover implementation details and basic exercises
 - Very important if you wish to get the maximum out of the course
- Topic list on the course schedule
- *Strongly recommend attending all recitations*
 - *Even if you think you know everything*

Evaluation

You are evaluated based on the following Tests:

- Weekly Quizzes
- Homeworks - 4
- Final Project (not for 11485)

Weekly Quiz

- 10 multiple-choice questions (3 attempts)
- Related to topics covered that week
 - On both slides and in lecture
 - Released Friday, closed Sunday night
 - This may occasionally shift, don't panic!
- There will be 14 total quizzes
 - We will consider the best 12
 - This is expected to account for any circumstance- based inability to work on quizzes
 - You could skip up to 2

Weekly Quiz

- Slides often contain a lot more information than is presented in class
- *Quizzes will contain questions from topics that are on the slides, but not presented in class*
- *Will also include topics covered in class, but not on online slides!*
- There will be questions based on latest research papers in the quiz – the links to the papers will be provided

Homework

- There will be one early homework (released before the start of the semester) and four in-term homeworks
 - Homework 0: Preparatory material for the course
 - Homeworks 1-4: Actual neural-net exercises
- Homeworks 1-4 all have two parts:
 - Part 1: Autograded problems with deterministic solutions
 - You must upload them to autolab
 - Part 2: Open problems posted on Kaggle
- Bonus HWs:
 - There will be 3 bonus homeworks corresponding to HW1-3 and 4 separate autograd bonus homeworks
 - These marks will not contribute to final grading curves but give you the chance to make up for marks missed elsewhere

Homework Part 1

- Part 1 of the homeworks evaluate your ability to code neural nets on your own from scratch
 - If you implement all mandatory and bonus questions of part 1 of all homeworks, you will, hopefully, have all components necessary to construct a little neural network toolkit of your own
 - “mytorch” 😊
- The homeworks are autograded
 - Be careful about following instructions carefully
 - The autograder is setup on a computer with specific versions of various packages
 - Your code must conform to their restrictions
 - If not the autograder will often fail and give you errors or 0 marks, even if your code is functional on your own computer
 - Note: Submission before ESD (early submission deadline) will get you an early submission bonus

Homework Part 2

- Part 2 of every homework tests your ability to solve complex problems on real-world data sets
 - The early submission deadline (ESD) is required and worth 10/100 points
 - The remaining 90/100 points are determined by your final score relative to the cutoffs

Homework part 2

- The standard cutoffs are (After 10pts ESD)
 - High cutoff : 90pts
 - Medium cutoff : 70pts
 - Low cutoff : 50pts
 - Very Low cutoff : 30pts
 - No submission : 0pts
- Your score is linearly interpolated between cutoffs. Ex, if an 85% accuracy model is for the high cutoff and 75% model is for the medium, a model with 80% accuracy will get 80/90pts
- There is a 5th bonus cutoff for extra credits (10pts) determined by the highest scoring student
 - Ex. If the highest accuracy is 92%, the scoring student will get 110/100 and anyone between 85% to 92% will get 100+/100 (considering they have the 10pts from ESD)

Homework Deadline

- Multiple deadlines
- Separate deadline for Autograded deterministic component
- Kaggle component has multiple deadlines
 - *Early Submission Deadline*: Worth 10% of P2 and affirms you have started the assignment
 - *On-time Submission Deadline*: Your final submission must occur before this deadline to be eligible for full marks (of your cutoff)
 - *Late/Slack Deadline*: Submissions after the On-time Submission Deadline will receive a penalty, and you can use slack days to avoid a penalty (but not to get bonus)
- Homeworks: Late policy
 - Everyone gets up to 10 total slack days (**does not apply to initial submission or part 1 of the HWs**)
- You can distribute them as you want across your HWs
 - You become ineligible for “A+” bonus if you’re using your grace days for Kaggle
- Once you use up your slack days, all subsequent late submissions will accrue a 10% penalty (on top of any other penalties)
- There will be no more submissions after the drop-dead deadline (Late/Slack)
- Kaggle: Kaggle leaderboards stop showing updates on full-submission deadline
 - But will continue to accept submissions until drop-dead deadline (using another Late kaggle)
- *Please see course webpage for complete set of policies*

Final Project

- If you're taking 11-785, you will be required to do a course project
 - 11-685 students will be assigned a fifth HW that is equivalent to a project
- Projects are done by teams of students
 - Ideal team size is 4
 - 11-685 teams are ideally 2-person teams
 - You are encouraged to form your teams early
- Projects are intended to exercise your ability to comprehend and implement ideas beyond those covered by the HWs
- Projects can range from
 - Implementing and evaluating cutting-edge ideas from recent papers
 - *Verifying* results from “hot” published work
 - “Researchy” problems that might lead to publication if completed well
 - Proposing new models/learning algorithms/techniques, with proper evaluation
 - Etc.

Final Project

- Think about forming project teams as soon as possible. More details on website.
 - If you don't form your own teams, we will team you up
- Each team must:
 - Submit a project proposal by the deadline listed on the website
 - Submit a mid-way report 3/4 way through the semester
 - Submit a preliminary full report three days **before the presentation due date**
 - Make a **5 min** video presentation of the project at the end of the semester
 - Can be presented by one, some, or all team members
 - Will be evaluated by the instructor, TAs, and your classmates
 - Ensure you explain the problem, proposed solution, and the evaluation clearly
 - Allocate enough time to make the presentation, it is not as easy as you think
 - Poor presentation can significantly affect your project score :)
 - Submit a final full report at the end of the semester
 - Defend your project in front of peers and TAs
 - Templates for proposals and reports will be posted
- *Each team will be assigned a mentor from among the TAs, who will monitor your progress and assist you if possible.*
- **More details on project evaluations will be posted towards the end of the sem**
- **The project is often the most fun portion of the course**

Grading

Weekly Quizzes **24**

14 Quizzes, bottom two dropped 24

Assignments **50**

HW1 – Basic MLPs	(AL + Kaggle)	12.5
HW2 – CNNs	(AL + Kaggle)	12.5
HW3 – RNNs	(AL + Kaggle)	12.5
HW4 – Sequence to Sequence Modelling	(AL + Kaggle)	12.5

Team Project (Not for 11-485) **25**

Proposal	-
Mid-term Report	5
Preliminary Full Report	-
Project Presentation	10
Peer Reviewing	binary multiplier
Final report	10

*** Note: There is 1 mark for attendance**

Preparation for the course

- Course is implementation heavy
 - A lot of coding and experimenting
 - Will work with some large datasets
- Language of choice: Python
- Toolkit of choice: Pytorch
 - You are welcome to use other languages/toolkits, but the TAs will not be able to help with coding/homework
 - Some support for TensorFlow
- We hope you have gone through
 - Recitation zero
 - HW zero
 - Carries no marks

Teamwork



- Learning happens best together
 - You will learn more from each other than you will from us
- *We encourage teamwork*
 - But there are strict rules...

Study Groups



- Please form study groups
- If you do not have a study group of your own, we will form one for you
 - Please register on the forms posted on Piazza
- *Everyone must be part of a study group*
- ***Each study group will be assigned a TA mentor for help throughout the course***

Study Group

- What study groups may do:
 - Discuss homework problems and solutions
 - Discuss papers
 - Discuss class work
 - Discuss quizzes
- We encourage you to meet regularly to discuss IDL work
- Study groups may also go on to form project teams
- So what are the caveats? What may you *not* do?

Study Group vs Cheating

- Every student must solve their quizzes by themselves
 - You may discuss the questions with your study groups/friends, but when you solve the quiz, **isolate yourself and do it alone**
- Every student must solve every homework by themselves
 - You may discuss the homeworks with your friends, and even help them debug their code, but when you finally solve it, every line of your code (except libraries that have been okayed by course staff) must be written by you
 - Your solution must be yours
- Plagiarizing code from the web or your friends constitutes cheating
 - And submitting solutions not obtained by you constitutes cheating

Cheating !!!

- You are here to learn DL yourself, not to demonstrate how well your friend, or that guy on the web has learned DL
- You are at CMU – which means you are among the brightest and best students in the *world*
 - You probably were among the top students in your peer group all your life, before you came here
 - It will be an insult to yourself and everything you ever stood for in your life to lower yourself from your own standards and start cheating
 - **So don't!!!**
- If you are unsure whether something you're doing constitutes cheating or not, check with us

Mentoring

- Every study group will be assigned a TA mentor
 - We will track your progress and reach out to you if you appear to be in trouble
- If in trouble, reach out to your TA mentor and/or the instructor
 - If you feel you're falling behind, reach out
 - If you feel you are struggling, reach out
 - If you feel pressured/unable to cope, reach out
 - We will try our best to help you
 - *Please watch Recitation 01 if you are stuck or feeling overwhelmed*
- We aim to make this a successful course for all of you
 - In our ID(ea)L world, everyone performs well enough to get an A
 - *Without lowering our standards* – i.e. we would like to bring you all up to where we believe you deserve an A
 - Everything about this course is geared to that objective

This Course is not easy !!!

- It's a lot of work !!
- It's a lot of work !!
- And, it's a lot of work !!



But.....



Do not hesitate to seek help whenever you need it. We are all here to help you.
The course is a lot of fun :)

What will you achieve...

- *Mastery-based* evaluation
 - Quizzes to test your understanding of topics covered in the lectures
 - HWs to teach you to implement complex networks
 - And optimize them to high degree
 - Projects to expose you to real world DL problems
 - **Target: Anyone who gets an “A” in the course is technically ready for a deep learning job**

Recitation 0

- Please, please, please, please, please go through the videos for recitation 0, and complete HW0.
 - These are essential for you to gain comfort with the coding require in the following homeworks
- HW1 part 1 also has many components intended to help you *later* in the course
 - So if it seems a bit dense, please bear with it, its worth it
- HW1 is the easiest HW!

Questions??

- Feel free to post on piazza for any clarification or doubts
- We strive to respond to student queries within 3 Mins



Thank you for your deciding to take this journey in Deep learning With us

Prof. Bhiksha Raj & TA's