

Course Project

Ju Sun

Computer Science & Engineering

University of Minnesota, Twin Cities

October 5, 2020

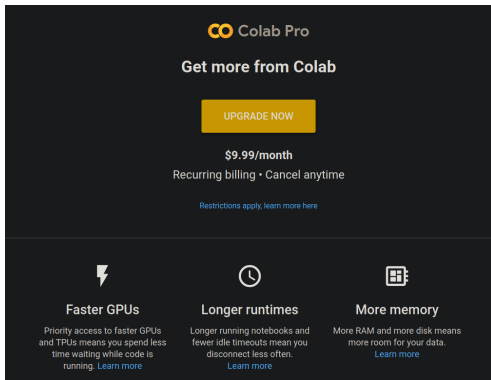
General logistics

Project logistics

Project ideas

Announcements

- HW1 is out. Due: Oct 15th
- Colab Pro <https://colab.research.google.com/signup>



The image is a promotional banner for Colab Pro. It features a dark background with white and yellow text. At the top, the Colab logo (two interlocking circles) is followed by 'Colab Pro'. Below this, the text 'Get more from Colab' is centered. A prominent yellow button with the text 'UPGRADE NOW' is in the center. Underneath the button, the price '\$9.99/month' is displayed, followed by 'Recurring billing • Cancel anytime'. A small link 'Restrictions apply, learn more here' is below that. The bottom section is divided into three columns, each with an icon, a title, and a description with a 'Learn more' link.




Colab Pro

Get more from Colab

UPGRADE NOW

\$9.99/month
Recurring billing • Cancel anytime

[Restrictions apply, learn more here](#)

		
Faster GPUs	Longer runtimes	More memory
Priority access to faster GPUs and TPUs means you spend less time waiting while code is running. Learn more	Longer running notebooks and fewer idle timeouts mean you disconnect less often. Learn more	More RAM and more disk means more room for your data. Learn more

General logistics

Project logistics

Project ideas

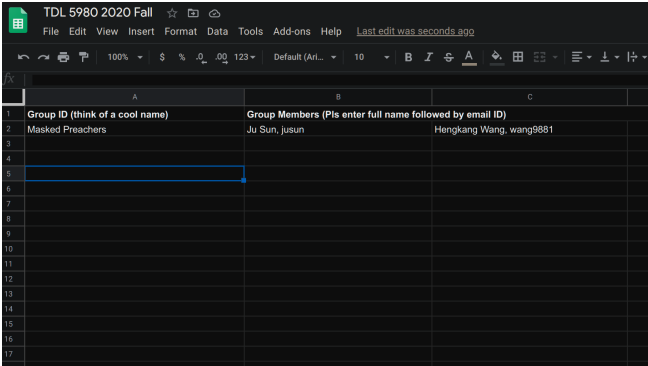
Timeline & L^AT_EX template

- Proposal (5%, 1 page): Oct 23
- Progress presentation (5%, 5 mins): Nov 23
- Progress report (10%, 3–4 pages): Nov 27
- Final report (20%, 7–8 pages): Dec 22
- Publishable results \implies A!

Template for all writeups: NeurIPS 2020 L^AT_EX style

<https://nips.cc/Conferences/2020/PaperInformation/StyleFiles>

Groups



The screenshot shows a Google Sheet interface with the title 'TDL 5980 2020 Fall'. The menu bar includes File, Edit, View, Insert, Format, Data, Tools, Add-ons, and Help. The toolbar shows various editing and formatting options. The spreadsheet has three columns: A, B, and C. Row 1 contains headers: 'Group ID (think of a cool name)' in column A, and 'Group Members (Pls enter full name followed by email ID)' in column B. Row 2 contains data: 'Masked Preachers' in column A, 'Ju Sun, jusun' in column B, and 'Hengkang Wang, wang9881' in column C. Rows 3 through 18 are empty, with row 5 highlighted in blue.

	A	B	C
1	Group ID (think of a cool name)	Group Members (Pls enter full name followed by email ID)	
2	Masked Preachers	Ju Sun, jusun	Hengkang Wang, wang9881
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			

- 5980: 2–3 students / 8980: 2 students
- All submissions as a group (in Canvas as group assignment); the group gets the same score

Proposal

- What problem?
- Why interesting?
- Previous work
- Your goal
- Plan and milestones

- Individual work (15%)
- Template for all writeups: NeurIPS 2020 L^AT_EX style
[https://nips.cc/Conferences/2020/
PaperInformation/StyleFiles](https://nips.cc/Conferences/2020/PaperInformation/StyleFiles)
- Due: Dec 05
- Talk to me about choice of topic

General logistics

Project logistics

Project ideas

Roughly by ascending level of difficulty

- Literature survey/review
- Novel applications
- Novel methods
- Novel theories

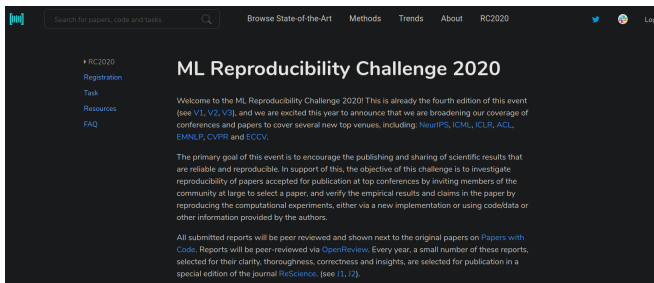
Excerpt from a research project is fine, but you should describe your own contributions

Literature survey/review

A coherent account of recent papers in a focused topic

- Description and comparison of main ideas, or
- Implementation and comparison of performance, or
- Both of the above

should **complement** the topics we cover in the course



<https://paperswithcode.com/rc2020>

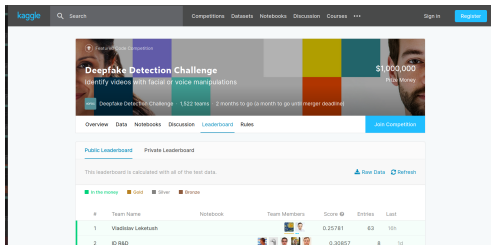
Random topics

- DL for noneuclidean data (e.g., graph NN, manifold NN)
- transformer models for sequential data
- generative models (e.g., GAN, VAE, normalization flow)
- 2nd order methods for deep learning
- differential programming
- universal approximation theorems
- DL for 3D reconstruction
- DL for video understanding and analysis
- DL for solving PDEs
- RL for games
- RL for robotics
- DL for medical imaging
- DL for (astro)physics
- DL for chemistry
- adversarial attacks; robustness of DL
- privacy, fairness in DL
- visualization for DNN
- network quantization and compression
- hardware/software platforms for DL
- automated ML; architecture search
- optimization/generalization theory of DL

Novel applications

Apply DL to **new** application problems

- A good place to start: Kaggle <https://www.kaggle.com/>



The screenshot shows the Kaggle website interface. At the top, there's a navigation bar with 'kaggle' logo, a search bar, and links for 'Competitions', 'Datasets', 'Notebooks', 'Discussion', and 'Courses'. A 'Sign In' button and a 'Register' button are on the right. The main content area features the 'Deepfake Detection Challenge' banner, which includes the text 'identify videos with facial or voice manipulations' and a prize of '\$1,000,000'. Below the banner, there are tabs for 'Overview', 'Data', 'Notebooks', 'Discussion', 'Leaderboard', and 'Rules'. The 'Leaderboard' tab is selected, showing a 'Public Leaderboard' and a 'Private Leaderboard' section. A note states 'This leaderboard is calculated with all of the test data.' Below this, there's a legend for 'In the money' (green), 'Gold' (yellow), 'Silver' (grey), and 'Bronze' (brown). A table lists the top teams:

#	Team Name	Notebook	Team Members	Score	Entries	Last
1	Vladislav Luketush			0.25781	63	10h
2	ID RAD			0.20857	8	1d

- Think about data availability

Google dataset search

<https://datasetsearch.research.google.com/>

- Think about GPUs

Where to find inspirations

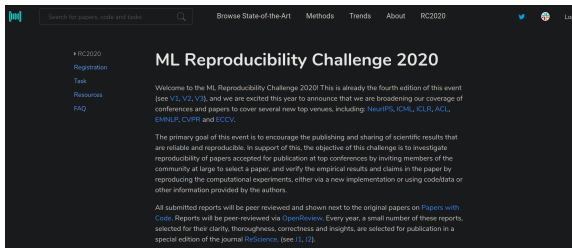
- arXiv machine learning
<https://arxiv.org/list/cs.LG/recent>
- Recent conference papers
 - ML: NeurIPS, ICML, ICLR, etc
 - CV: ICCV, ECCV, CVPR, etc
 - NLP: ACL, EMNLP, etc
 - Robotics: ICRA, etc
 - Graphics: SIGGRAPH, etc
- Talk to researchers!

Novel methods

Create new **NN models or training algorithms** to improve the state-of-the-art

Where to start:

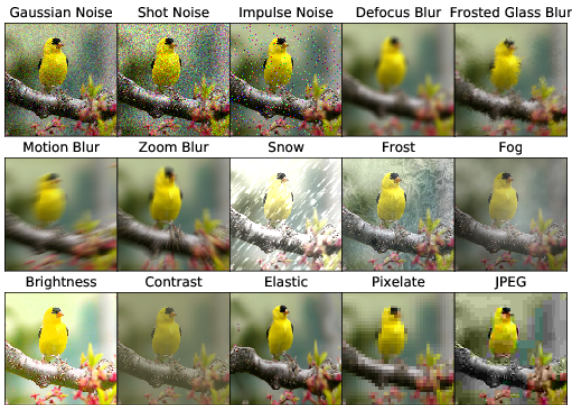
- Kaggle (again)!
- arXiv machine learning and recent conference papers
- MLRC



<https://paperswithcode.com/rc2020>

Novel methods

Equally interesting to fool/fail the state-of-the-art, i.e., exploring robustness of DL



Credit: ImageNet-C

Novel theories

Nothing is more practical than a good theory. – V. Vapnik

- universal approximation theorems
- nonconvex optimization
- generalization

Where to start:

- Analyses of Deep Learning (Stanford, fall 2019)
<https://stats385.github.io/>
- Theories of Deep Learning (Stanford, fall 2017)
https://stats385.github.io/stats385_2017.github.io/
- Toward theoretical understanding of deep learning (ICML 2018 Tutorial)
<https://unsupervised.cs.princeton.edu/deeplearningtutorial.html>

Questions?