ECE 1786:
Creative Applications of Natural Language Processing

Fall 2022

Instructor: Jonathan Rose

Department of Electrical & Computer Engineering

Land Acknowledgement

Toronto is a city of immigrants. My parents came here in 1952, and it is very likely that you or your parents are very new to this country. This is not a new feature of Toronto. For over 15,000 years Toronto has been a gathering site for humans including the Huron-Wendat and Petun First Nations, the Seneca, and most recently, the Mississaugas of the Credit River.

Today, Toronto is still a meeting place for Indigenous people from across Turtle Island (North America), and immigrants, both new and old, from across the world. I am grateful to have the opportunity to work in this community, and on this territory.

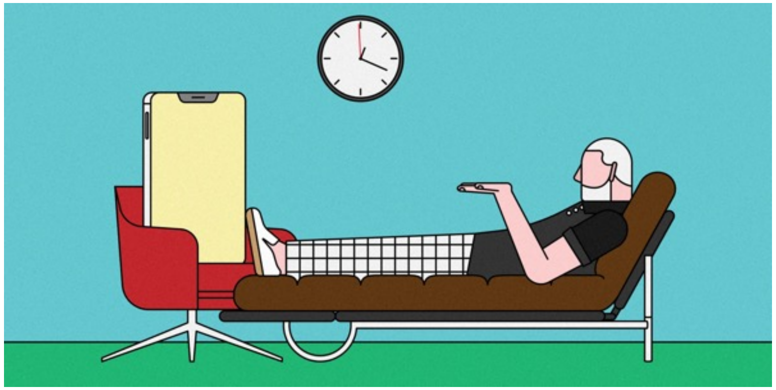
Welcome to ECE 1786!

- This course is about methods for automatic understanding and generation of human language
- Language is a pathway in and out of the brain!
- Until recently, it was difficult to connect humans and computers with language
- Course covers inventions that have made sophisticated communication possible and has led to other remarkable capabilities

Those Capabilities Include …

- In your lifetime, you will be able to talk to a computer-based doctor to get decent basic medical care

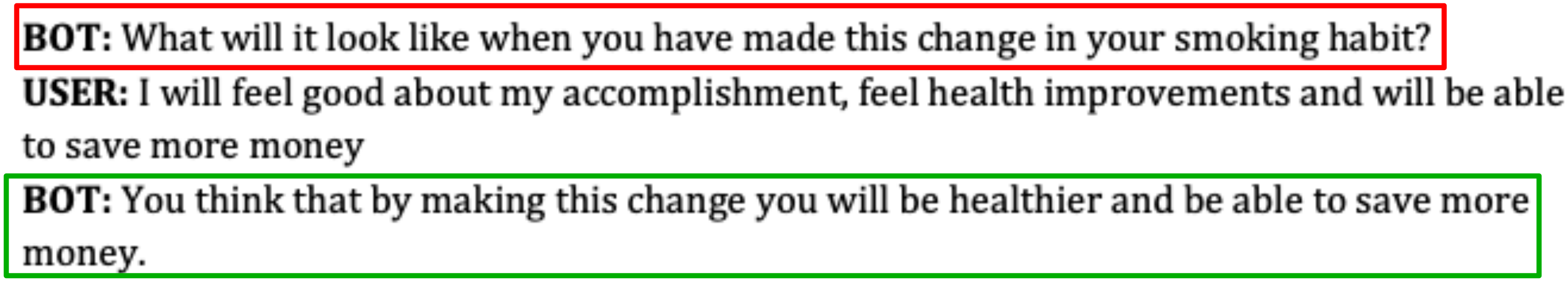
- You'll be able to get mental-health *talk therapy* any time, anywhere from a computer.



My Current Research Focus

- Automatic Talk Therapy for Smoking Addiction
  - As one example of talk therapy

Scripted



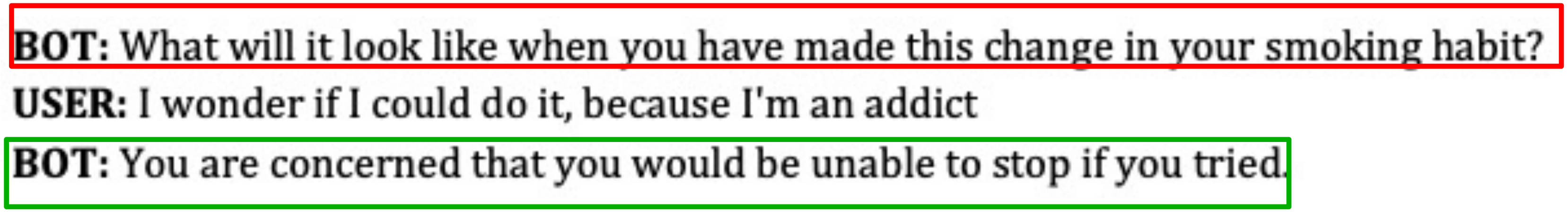
BOT: What will it look like when you have made this change in your smoking habit?

USER: I will feel good about my accomplishment, feel health improvements and will be able to save more money

BOT: You think that by making this change you will be healthier and be able to save more money.

Generated!

Scripted



BOT: What will it look like when you have made this change in your smoking habit?

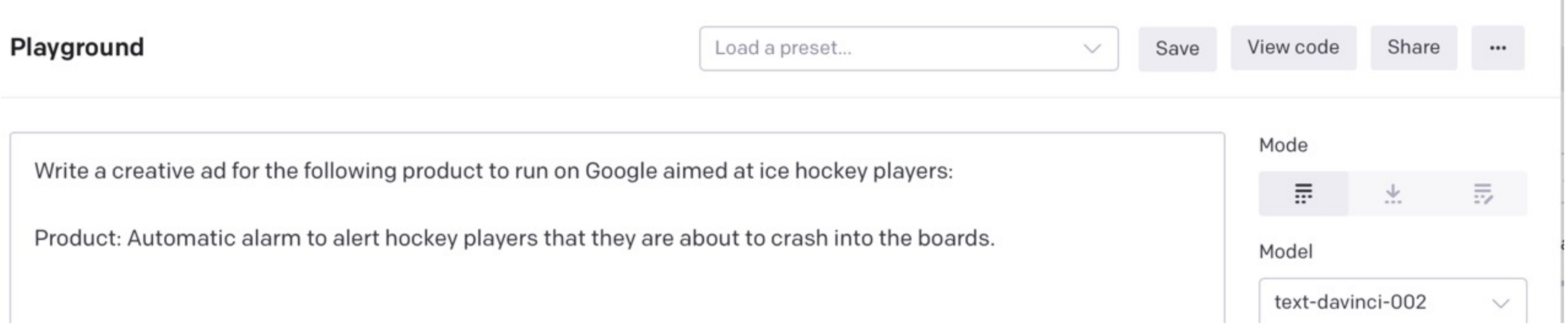
USER: I wonder if I could do it, because I'm an addict

BOT: You are concerned that you would be unable to stop if you tried.

Generated!

GPT-3 Playground

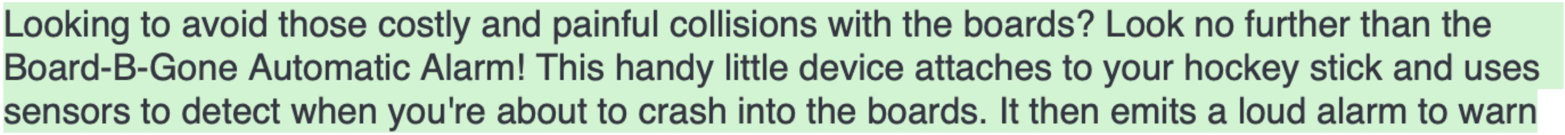
- State-of-the-art language capability at your fingertips!
  - https://beta.openai.com/playground



Write a creative ad for the following product to run on Google aimed at ice hockey players:

Product: Automatic alarm to alert hockey players that they are about to crash into the boards.

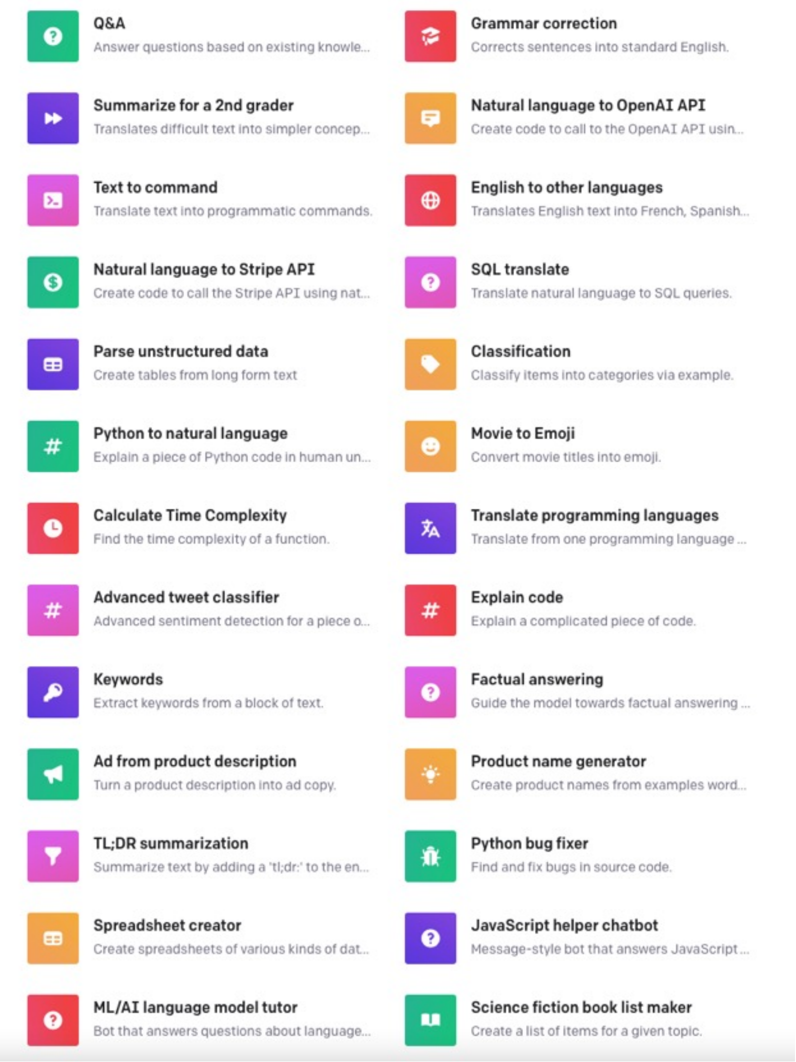
Model
text-davinci-002



Looking to avoid those costly and painful collisions with the boards? Look no further than the Board-B-Gone Automatic Alarm! This handy little device attaches to your hockey stick and uses sensors to detect when you're about to crash into the boards. It then emits a loud alarm to warn

More Examples from GPT-3

- See https://beta.openai.com/examples



Q&A
Answer questions based on existing knowle...

Summarize for a 2nd grader
Translates difficult text into simpler concep...

Text to command
Translate text into programmatic commands.

Natural language to Stripe API
Create code to call the Stripe API using nat...

Parse unstructured data
Create tables from long form text

Python to natural language
Explain a piece of Python code in human un...

Calculate Time Complexity
Find the time complexity of a function.

Advanced tweet classifier
Advanced sentiment detection for a piece o...

Keywords
Extract keywords from a block of text.

Ad from product description
Turn a product description into ad copy.

TL;DR summarization
Summarize text by adding a 'tl;dr:' to the en...

Spreadsheet creator
Create spreadsheets of various kinds of dat...

ML/AI language model tutor
Bot that answers questions about language...

Grammar correction
Corrects sentences into standard English.

Natural language to OpenAI API
Create code to call to the OpenAI API usin...

English to other languages
Translates English text into French, Spanish...

SQL translate
Translate natural language to SQL queries.

Classification
Classify items into categories via example.

Movie to Emoji
Convert movie titles into emoji.

Translate programming languages
Translate from one programming language ...

Explain code
Explain a complicated piece of code.

Factual answering
Guide the model towards factual answering ...

Product name generator
Create product names from examples word...

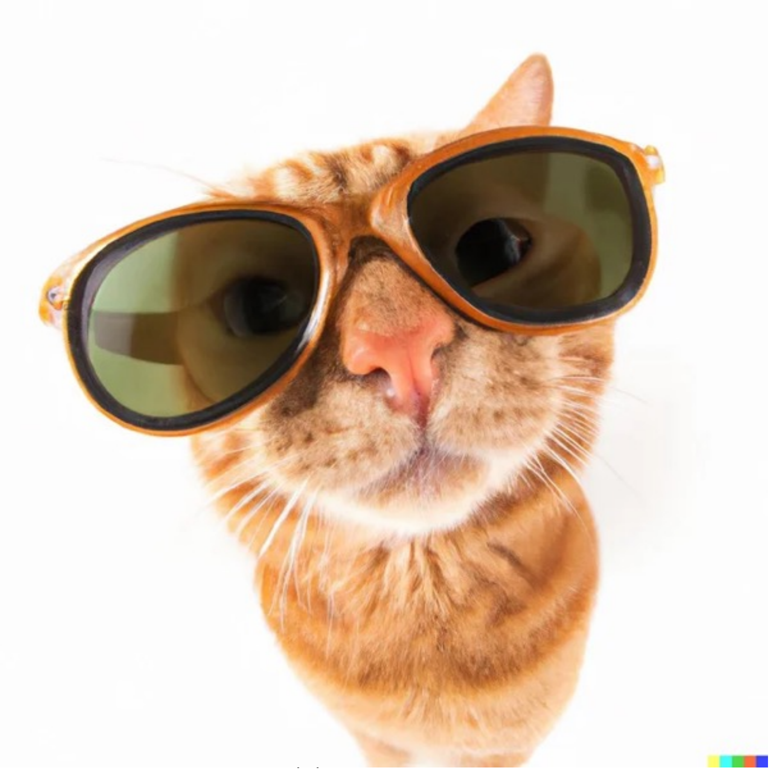
Python bug fixer
Find and fix bugs in source code.

JavaScript helper chatbot
Message-style bot that answers JavaScript ...

Science fiction book list maker
Create a list of items for a given topic.

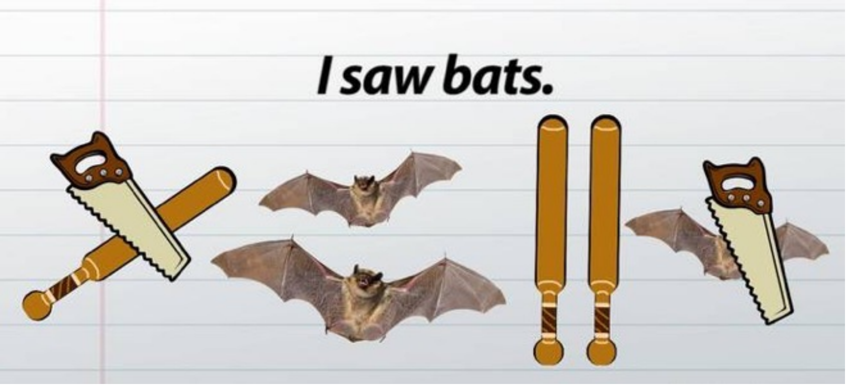
NLP Advances are Part of DALL-E-2

- One of several amazing models that turn can turn words into remarkable pictures:
  - “A closeup wide angle photo of a cat wearing sunglasses”
  - See https://www.reddit.com/r/dalle2/ for more examples



What Makes Language Difficult

- The ambiguity of language:
  1. Multiple meanings of words
     - tank, bank, duck ….
  2. Context needed to figure out:
     - Milk drinkers are turning to powder
     - Juvenile Court Tries Shooting Defendant
     - Grandmother of Eight makes Hole in One

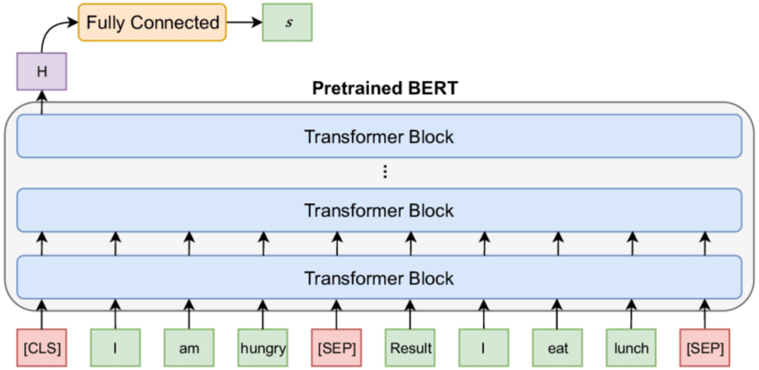


Is Now Much Easier to Deal with

- Using neural-net-based approaches to NLP
- Since the Deep Learning revolution began in 2012, many fields have been changed
  - A key step in word embeddings occurred then
- Key next steps for NLP occurred in 2018
  - ‘Transformer Architecture’ Vaswani et. al
  - BERT - https://huggingface.co/bert-base-uncased
  - GPT-2 - https://huggingface.co/gpt2
  - GPT-3 - https://openai.com/blog/gpt-3-apps/
  - Bloom - https://huggingface.co/bigscience/bloom
  - LaMDA - https://blog.google/technology/ai/lamda/

Success in Key Applications

1. Translation
2. Classification
3. Dialogue



Related Courses & Focus of this Course

Course Focus vs. Related Courses

- Course focus is neural-network approaches to NLP
- NLP is a 60-year-old field which previously used a procedural methodology:
  - Based on human understanding of language
  - Referred to as ‘Computational Linguistics’
    - Grammar – e.g. parts of speech tagging – noun, verb
    - Parsing of language to interpret
  - CSC 485/2501 - Title: **Computational Linguistics**
  - https://www.cs.toronto.edu/~gpenn/csc485/
  - Principally taught by Professor Gerald Penn
- We will make some, limited use of computational linguistic approach

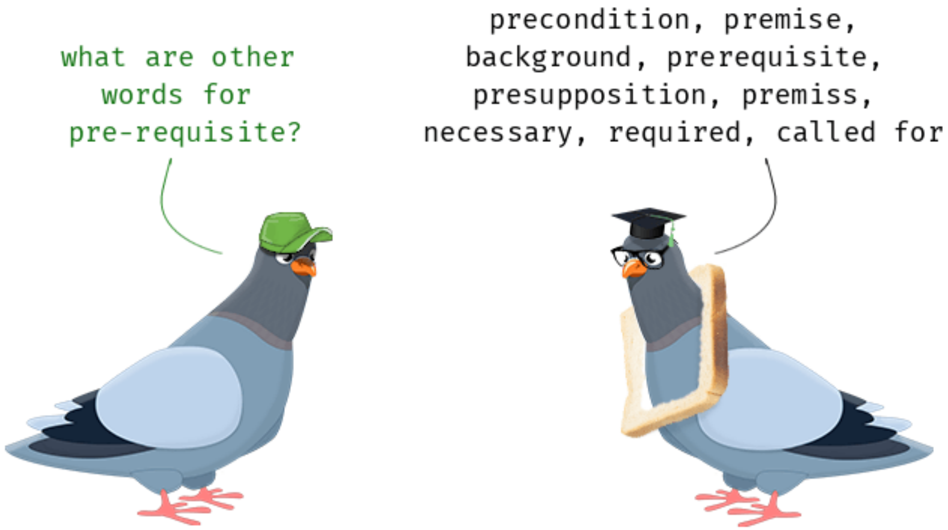
Course Focus vs. Related Courses

- A more closely related course in Computer Science
- CSC 401/2511 - **Natural Language Computing**
  - https://www.cs.toronto.edu/~frank/csc401/
  - Principally taught by Professor Frank Rudzicz
- CSC 2511 has a broader coverage of topics surrounding language
  - a super-set of this course,
  - including things such as Markov Models, Entropy, Automatic Speech Recognition, retrieval and dialogue
- ECE 1786 focuses more narrowly on word embeddings, deep learning, statistical language models,Transformers
  - more depth less breadth
  - more of an engineering, software focus

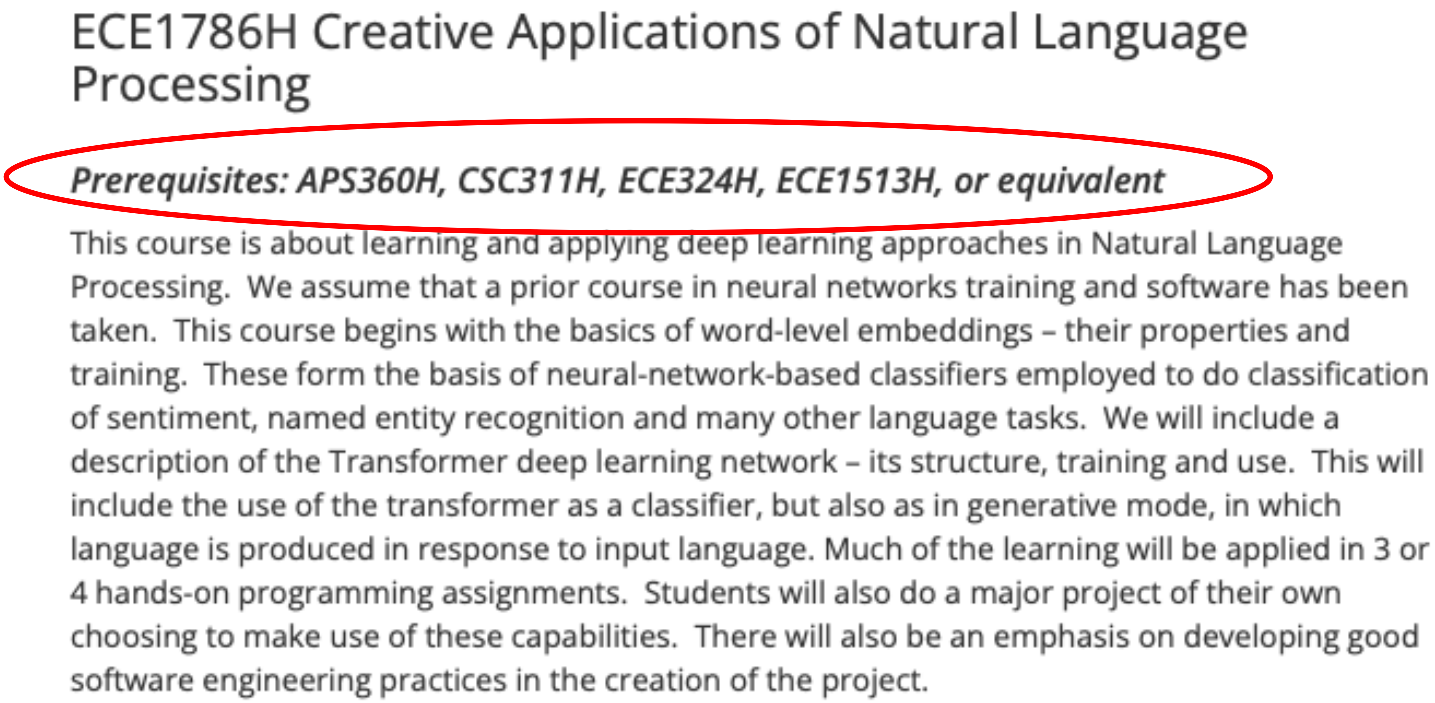
Learning Outcomes: Understanding

- Word embeddings
- Use of word embeddings in classification tasks
- Transformers
  - Global structure
  - Training
  - Attention, Encoder-Decoder, Transformer Stack
  - Classification
  - Probabilistic & Auto-regressive Generation
- PyTorch, Huggingface
- Limitations and Biases of Language models
- Navigation of open-ended problems in a project

Course Pre-Requisites


what are other words for pre-requisite?
precondition, premise, background, prerequisite, presupposition, premiss, necessary, required, called for

We Need To Talk about Pre-requisites



ECE1786H Creative Applications of Natural Language Processing

Prerequisites: APS360H, CSC311H, ECE324H, ECE1513H, or equivalent

This course is about learning and applying deep learning approaches in Natural Language Processing. We assume that a prior course in neural networks training and software has been taken. This course begins with the basics of word-level embeddings – their properties and training. These form the basis of neural-network-based classifiers employed to do classification of sentiment, named entity recognition and many other language tasks. We will include a description of the Transformer deep learning network – its structure, training and use. This will include the use of the transformer as a classifier, but also as in generative mode, in which language is produced in response to input language. Much of the learning will be applied in 3 or 4 hands-on programming assignments. Students will also do a major project of their own choosing to make use of these capabilities. There will also be an emphasis on developing good software engineering practices in the creation of the project.

- Graduate office does not check pre-requisites
- I will do that now, in two ways …

This Course does not introduce ML

- You **must** have background from a course in machine learning that has depth in neural networks
  - Otherwise, you won't understand what to do in the assignments!

- Acceptable UofT undergraduate courses (any one)
  - ECE 324 Machine Intelligence, Software and Neural Networks
  - ECE 421 Introduction to Machine Learning
  - APS 360 Applied Fundamentals of Deep Learning
  - CSC 311 – Introduction to Machine Learning
  - CSC 413 – Neural Networks and Deep Learning

- ECE 1513 Introduction to Machine Learning

Pre-requisite, cont’d

- If undergraduate degree is from elsewhere (most of you):
  - You must have taken a course that is equivalent to one of these University of Toronto courses
  - Next few slides describe the necessary background in detail

- Everyone must fill out the Quercus survey, posted on the course main website,
  - You’ll say which course you’ve taken, and provide a link to any non-UofT course
  - You’ll try to answer questions based on the following….

What to Know Already

- Machine Learning
  - Classification vs. regression (Logistic v. Linear regression)
  - Binary vs. multi-class classification
  - Supervised vs. unsupervised learning
  - Data labelling

What to Know Already

- Basic Neural nets and training
  - Linear neurons – weights and biases
  - Non-linear activation functions, e.g sigmoid, ReLU …
  - Multi-layer perceptron (MLP)
  - Loss functions; binary & multinomial cross entropy
  - Softmax function
  - Training, Validation and Test sets
  - Training & Validation ‘curves’
  - Gradient Descent
  - Stochastic Gradient Descent
  - Hyper-parameter tuning
  - Regularization: normalization, dropout, weight decay

What to Know Already, continued

- Advanced Neural Networks
  - Convolutional Neural Networks for computer vision (CNN)
    - Kernels, batch normalization
  - Recurrent Neural Networks (RNN)

- Transfer Learning:
  - Pre-trained networks
  - Fine-tuning of pre-trained networks

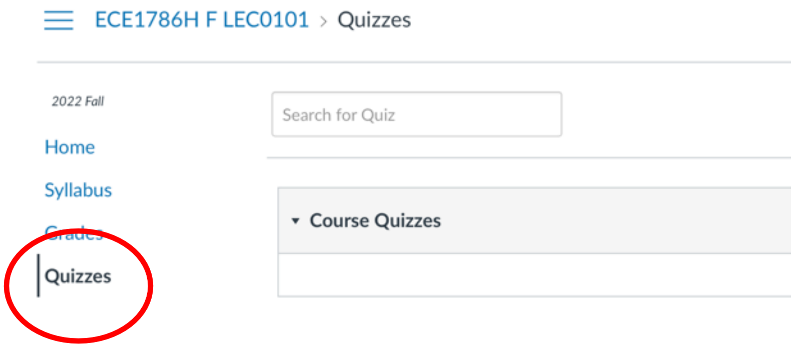
What to Know Already, continued

- **How to build all that** with Software Frameworks
  - Experience in Tensorflow or PyTorch (implies Python experience)
  - Should have written full training and test loops applications with significant data sizes
  - Tensors, shape
  - Numpy framework,
  - **How to debug a neural network**

- Data Science concepts
  - False positive, False Negative, True Positive, True Negative
  - Sensitivity, Specificity
  - Area Under the ROC Curve
  - Confusion Matrix

For you to do after finishing this video

- Go to the Quercus Website for this course
  - Assuming you're enrolled and not on the waitlist
- Click on 'Quizzes' on the left-hand side
  - Answer all of the questions to the best of your ability
  - This will not be graded
  - You must give evidence that you have the pre-requisite knowledge and training to remain in course



ECE1786H F LEC0101 > Quizzes

2022 Fall

Teaching Philosophy

Teaching Philosophy

1. Teach less, but more in-depth
   - try to explain the intuition behind each approach
   - believe with solid grounding, you can learn more on your own

2. Your learning comes from doing.
   - From writing or working with software, experimenting with it
   - Answer questions about the results of experiments

3. Do open-ended projects
   - An engineer can navigate an open-ended project
   - Learn to conceive a project & describe it: **what** & **why**
   - Then **do it** – figure out **how**, and make it happen
   - Then **describe it**

Course Structure & Grading

Grading

Item	Fraction of Course
Assignments (4)	40%
Project Proposal Document/Presentation	10%
Project Interim Report	10%
Project In-Class Final Presentation	10%
Project Peer Reviews	5%
Project Final Report/Software	25%
Total	**100%**

Project is 60%

Assignments

#	Date Assigned*	Assignment	Due*
1	September 13	Word Embeddings – Properties, Meaning and Training	September 26th
2	September 27	Classification of Subjective/Objective Text	October 10
3	October 11	Understanding, Training and Using Transformers	October 24
4	October 25	Question Answering Using Transformers	November 14

*Tentative dates, subject to change

Peer Review

- Each individual student will be asked to provide feedback to other groups on their:
  - Proposal presentation/document
  - Interim report
- Asked for specific/useful feedback to group’s work
- Feedback/commentary will be graded for quality

Textbooks & Hardware Acceleration

Required Text is Free: **Speech and Language Processing (3rd Edition Draft)** by Dan Jurafsky and James H. Martin:

https://web.stanford.edu/~jurafsky/slp3/ed3book_jan122022.pdf

2nd Edition has a complete first chapter (missing above):

https://github.com/rain1024/slp2-pdf

For the assignments and project, it is suggested that you purchase the for-pay Google Colab Pro:

https://colab.research.google.com/signup

The Project

The Project

- Done in Groups of 2
  - No groups of 1 or 3
- The topic is of your own choosing
  - must be approved by instructor
- Must relate to Natural Language Processing & the material covered in this course
- It should be an application of NLP
  - Need permission to do a research project on NLP itself
- Must collect and label some of your own data
  - Cannot simply work on an existing dataset
  - Why? Data labelling is at the core of all ML/AI work
  - However, this can must be careful not to do too much!

Project Stages

1. **Forming Groups**
   - Should be done by end of October
2. **Project Approval-in-Principle**
   - via email; due October 24th
3. **Project Proposal/Plan**
   - Document Due October 31
4. **Proposal & Plan Presentations**
   - November 1
   - **NOTE EXTRA LECTURE Tuesday November 1st, 6-8pm**
5. **Interim Report**
   - November 21
6. **Final Presentations**
   - December 6th; extra lecture that week as well.
7. **Final Report Due December 7**

- All dates are tentative

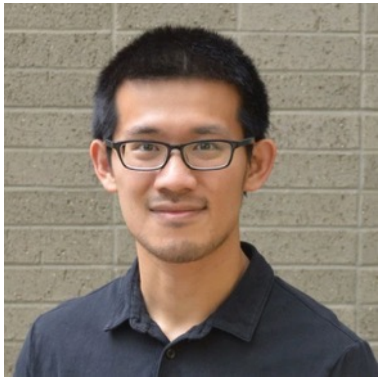
Course Instructor/TAs

Instructor Bio: Jonathan Rose

- Professor in Electrical & Computer Eng since 1989
  - Bach, Master's & PhD from UofT, Post-Doc at Stanford
- Research: Automation of Medicine/Mental Health
  - Automation of Mental Health using Machine Learning/NLP
    - Focusing on conversational systems for mental health
  - Previously: Field-Programmable Gate Arrays (FPGAs)
- Entrepreneurial/Business Experience:
  - Co-founder of Right Track CAD Corp in 1998
  - Software Engineering Director of Altera 2000-2003, now Intel
- Administration:
  - ECE Dept. Chair of ECE 2004-2009;
  - Chair Engineering Entrepreneurship **Hatchery** Advisory Board
- F.IEEE, F.ACM, F.CAE, FA NAE, FRSC, Sr Fellow Massey College

Teaching Assistants

- Zining Zhu
  - Ph.D. Candidate in Computer Science
  - Thesis: Method and Applications for Probing Deep Language Models
  - Zining.zhu@cs.toronto.edu



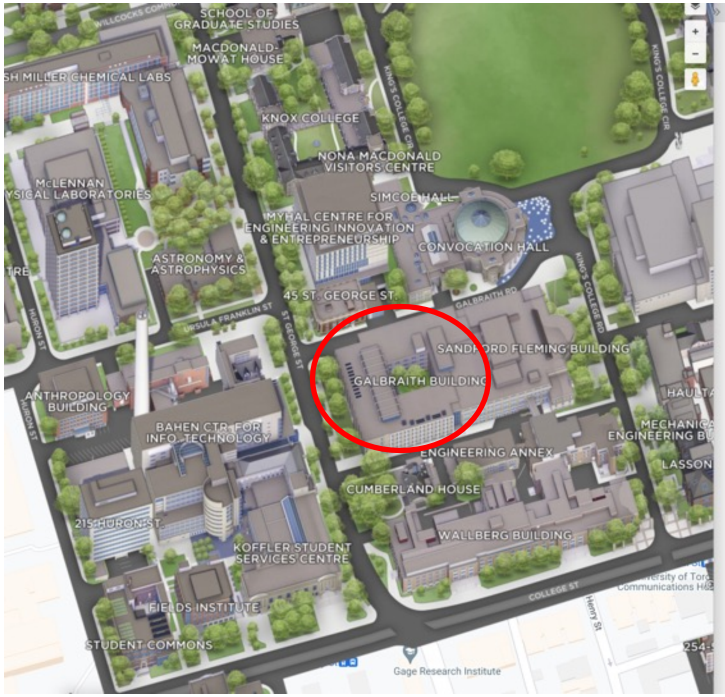
- Andrew Brown
  - M.A.Sc. Candidate in ECE
  - Thesis: Using NLP for a Smoking Cessation Chatbot
  - andrewm.brown@mail.utoronto.ca

Two Course Websites:

- UofT Quercus (https://q.utoronto.ca/courses/285821) for
  - Assignments release and submitted
  - Grades
  - Announcements
- **Piazza** website for a discussion board
  - See announcement on Quercus that tells you how to access
  - Email me if you don’t have access to Quercus & I will add you

Questions?

- Post them to the Piazza discussion board for this course
- **or**, bring them to the first lecture on
  - Day: September 13th, 2022
  - Time: 10am-12 noon
  - Place: Galbraith Building, 35 St. George Street, Room 221:
  - See you there then!


SCHOOL OF GRADUATE STUDIES
MACDONALD-MOWAT HOUSE
SH MILLER CHEMICAL LABS
KNOX COLLEGE
KING'S COLLEGE CIR
NONA MACDONALD VISITORS CENTRE
McLENNAN YSICAL LABORATORIES
SIMCOE HALL
MYHAL CENTRE FOR ENGINEERING INNOVATION & ENTREPRENEURSHIP
CONVOCATION HALL
ASTRONOMY & ASTROPHYSICS
45 ST. GEORGE ST
GALBRAITH RD
URSULA FRANKLIN ST
ST GEORGE ST
HURON ST
SANDFORD FLEMING BUILDING
GALBRAITH BUILDING
ANTHROPOLOGY BUILDING
BAHEN CTR FOR INFO. TECHNOLOGY
HAULTA
MECHANICAL ENGINEERING BU
ENGINEERING ANNEX
LASSON
CUMBERLAND HOUSE
215 HURON ST
WALLBERG BUILDING
KOFFLER STUDENT SERVICES CENTRE
FIELDS INSTITUTE
STUDENT COMMONS
COLLEGE ST
Henry St
Gage Research Institute