## WELCOME



## WHAT YOU'LL LEARN IN THIS COURSE

# NUMERICAL ALGORITHMS PYTHON DATA ANALYSIS

### NUMERICAL ALGORITHMS

- 1. LU Decomposition
- 2. Interpolation and Extrapolation
- 3. Solving integrals
- 4. Solving differential equations
- 5. Grid based methods
- 6. Monte Carlo Sampling

### PYTHON

- 1. Simple programs
- 2. Floating point numbers
- 3. Working with arrays and matrices
- 4. Plotting

#### DATA ANALYSIS

- 1. Working with python
- 2. Jupyter notebooks workflow
- 3. Real data, collected by yourself
- 4. Monte Carlo

#### WHY?





- Knowledge about machine learning the , and algorithms: SVM, random forest, gradient boosting methods, graphical models, bayesian methods, etc.
- Some Knowledge about deep learning theory, algorithms and tricks: RNN, ConvNet, seq2seq, dropout, attention mechanism, data augmentation, batch normalization, adversarial learning, VAE, GAN, etc.
- Knowledge about model optimization methods: 1st/2nd order methods, bayesian optimization, etc
- · Algorithm, data structure and object oriented programming skills
- Proficiency in Python
- Experience in applying machine learning algorithms for natural language processing/generation tasks
- Experience with the Linux stack(bash, git, package management etc.)
- Experience in processing large amounts of data

## SYLLABUS

#### GENERAL

- Office hours: Monday 11 - noon (today: 1pm-2pm) Tuesday 1pm - 2pm Any other time!
- hanno.rein@utoronto.ca
- @hannorein

#### LECTURES

- Be on time!
- 10 minute break
- No food in the lecture
- Notes on paper are encouraged
- No phones
- Computers allowed, but only if course related, no Facebook

#### TUTORIALS

- Be on time!
- Bring your computer! If you use the lab computers, bring your USB stick.
- Python introduction, help with assignments, ask questions about the course, quizzes

### **ASSIGNMENTS / QUIZZES**

- Coding assignments
- Submit the jupyter notebook (ipynb)
- Quizzes in lecture or tutorial
- Quizzes will test your knowledge about
  - Course material
  - Assignment
  - YOUR assignment

## PLAGIARISM

## OK

- Using the internet
- Asking your professor / TA for help
- Talking to other students

## NOT OK

- Copy and pasting code from the internet or other students
- Not understanding what you submit



## You have to understand what you submit.

#### GRADES

Assignments and quizzes	25*
Geotab project	20
Midterm	20
Final exam	35

\*Need at least 40% in assignments/quizzes to pass the course

## **GEOTAB PROJECT**

### WHAT IS IT?

- Team based project you'll work on the entire term
- Devices generously provided by Geotab
- 20% of your grade
- 3 deliverables: proposal / report / presentation

### WHAT DO YOU HAVE TO DO?

- Use a GPS tracker for cars to collect data
- Come up with a clever idea on how to use the data
- Write a program in python



#### **MYGEOTAB**



#### Fuel Usage Reports



#### Fleet Savings Report



#### **Driver Safety Scorecard**



#### **Risk & Safety Reports**





### FIRST GEOTAB DELIVERABLE IS DUE OCT 20TH!

- Start to think about project ideas now!
- Start to form teams (5-6 students)!
- Try to form a team with diverse skills! Not just your friends!
- Team formed? Come see me to get a geotab device!

## ON CHOOSING THE RIGHT PROJECT IDEAS

- Don't try to do something that is too complicated!
- Do something simple and focus on the report, the execution, and the implementation.
- If possible, try to make it *scientific*.

### WHY THIS GEOTAB PROJECT?

- Real world data
- Apply what you learned:
  - Python, jupyter-notebooks
  - Working with lists and arrays
  - Interpolation, extrapolation
  - Plotting
- How to solve a problem when something doesn't work

## SOFTWARE

#### SOFTWARE

You need access to a computer with

- Python 3.x
- Jupyter-notebooks
- numpy, scipy, matplotlib

#### SOFTWARE

Two options:

- Install it on your own computer (e.g. anaconda)
- Boot linux from a USB stick and use any computer, including the lab computers in physics

## FLOATING POINT NUMBERS

