HU Extension School E-89 Big Data Analytics

Handed out: 04/16/2020 **Due by 09:00 PM EST on Wednesday, 05/13/2020**

Final Project

You must select a technical topic related to Deep Learning from the list of available topics or suggest your own topic. Your final project must meet requirements as outlined below provide a not-trivial demonstration of a selected technology or a selected use case.

This final project is an individual effort.

The purpose of the final project is for each one of us to learn a new Deep Learning technology, or use case, not covered in class and then our colleagues gain the same knowledge. Write your final project report as a tutorial or a learning tool. All project materials including all code will be made available to the class. Please do not include any proprietary code or code you do not want to become public. This does not mean that you can do a project and then refuse to share your code. No working code, no grade. Please deliver your projects using Keras, TensorFlow or PyTorch APIs.

Topics proposed by the teaching staff will be distributed in an on-line auction. The auction allows 2 individuals to sign up for each topic. The auction is on the first come, first served basis. For the auction, we will use an online tool called Signup Genius. In a separate email, you will receive the link to the tool.

You may propose your own topic. Such topic must be approved by the professor**.** Reuse of topics and materials from previous classes is not acceptable.

You will receive an email with login to Signup Genius web site where you will be able to select your topic. Expect to get an email from by April 19th with an invitation to sign up for a topic. By April 18th the teaching stuff will post a list of possible topics and send the invitation email. The invitation email will be sent to the email address you registered with the school for this class. If you do not get an email, please contact us through Canvas.

Teaching staff does not provide guidance on the final project implementation. You must do the research yourself. Once you sign up for a topic, the topic is yours.

The AUCTION WILL START on Sunday, April 19th, 2020 at 09:00 AM EST.

Your project must demonstrate a Deep Learning related method, algorithm or a technology or a specific use case and include working software demonstration. Your project must include a visualization component with graphical presentation of your

results. The final project is not a research paper but rather a technology demonstration and a tutorial.

All students will present their short YouTube videos **between 12:00 (noon) PM and 3:00 PM on Saturday, May 16th. Please note extended hours.** All students are expected to participate through ZOOM and answer questions from the audience. **Your short YouTube presentation will be run on instructor's machine.** Every short YouTube video is expected **to last not** (much) **more than 2 minutes**. Q&As should not last more than a minute. **Note: if you are not present during airing of your video and the subsequent Q&A, you will not receive full credit for your presentation.**

Final Project Requirements:

1. Select a technology topic from the auction list via Signup Genius starting on Sunday, April 19th, 09:00 AM EST.
2. Write/define a problem statement describing what you will solve/resolve/demonstrate using the selected technology.
3. Select a data source (there are many data sets available online).
4. Develop/implement/code a solution for the specified problem using selected data source and selected technology using Keras, TensorFlow or PyTorch APIs
5. Produce a visualization of your data/results.

Documentation Requirements:

a) **Project summary (Abstract), 1 (one) page Word document in a separate file. Project summary should contain your name and the title of your project and URLs to 2 YouTube videos.**

   Produce a one page summary to describe the problem you are trying to solve and demonstrate, your data set, the particular technology or feature you are demonstrating, its uses, benefits, drawbacks, challenges and your results. Describe briefly the working example you prepared. This first page is very important. Based on that page, your colleagues will judge whether to download your project and spend time reading your documentation and code. Please add both YouTube URLs to the bottom of the one page summary.

   Note: The filename should start with the topic name followed by your name.

b) **Demo/solution implementation and working code:**

   Produce a single working demo that meets your problem statement and provides a full implementation of your solution. You may not simply copy examples (demonstrations) of the technology that are on-line. (Grade will be 0). You can use code skeletons that are provided as we have done in class and show your extended

programming use of them. **Provide neatly organized and complete working code with comments**. Please note that a project will not be given any points if a working implementation and code are missing. Please provide the URL to your full data source. If your data set is larger than 20MB, PLEASE DO NOT UPLOAD your data set. In such a case, please provide only a sample of your data with the rest of your submission. Our site has a limited capacity and a few large submissions could block it.

PLEASE, PLEASE do not upload final projects which take more than 30 MB of space. In the past, we had students who would upload GBs of data and block Canvas site. We do not find such practical jokes amusing.

c) **Slides:**

Produce a set of Power Point slides (10-20) with a few snapshots of your demo which captures the key points: your problem statement, what the technology does, your demonstration and pros/cons. The first page of your Power Point slide must have a standard format that we provide. Please use white background for all your Power Point slides. White background makes slides readable, printable and presentable on YouTube. Place URLs of your YouTube videos on the last slide. Note: The filename should start with the topic name followed by your name

d) **Report:**

Produce a detailed document with a complete description of analyzed technology or use case including all installation and configuration steps. Your report will start with your name and project title. **Detailed installations and configuration is required. Your colleagues must be able to reproduce your work based solely on the steps you have documented.** Describe your problem statement, data set, installation/configuration, results, what worked, what did not and why not, and any lessons learned. Report must show all steps to reproduce examples that you developed. This report is similar to your homework solutions where all steps are described along with the results.

a) The first page of your report should be the same as the one page summary (abstract). The first page must contain your name and the project name.
b) Please include your name and your project topic in the header or the footer of your MS Word Document.
c) Name your file like you named 1 page summary appended with _report.
d) Please include page numbers.
e) You are welcome to upload a PDF version of all documents. However, you must always submit an MS Word version of your report. Just submitting a Jupyter notebook is not acceptable. You must create a written report in MS Word.

f) Please include URLs for both short and long YouTube videos at the end of your report

e) **15 minute YouTube presentation:**

Produce a 15 minute YouTube presentation. This video will contain a summary of the technology and details of your project and results. If you insist, you can record two or even three15 minute presentations. Please do not record four.

f) **2 minute YouTube presentation:**

Produce a 2 minute YouTube presentation. This video will contain a summary of the technology and a quick overview of your demonstration. This video will be presented to the entire class so please make sure it is 2 minutes and not (much) more.

Submission Instructions:

Submit final project on the course site. **Upload your reports and slides separately.** Upload your zip file that contains your software and data files separately. Include your visualization files: html, JavaScript, Python, R code and data file within a separate folder in the zip file. Ensure your visualization files and all dependent files run! TAs will deduct points if they have to edit your code to run it. Name your zip file: **ShortProjectName_YourLastNameYourFirstName_Final.zip**

Grading criteria:

If you fail to provide working code submitted in working directories with all artifacts it produces and (toy-sample) dataset, your project will not be graded.

Code and surrounding artifacts are submitted in a ZIP file.

All other artifacts: Report, One page summary, PPTs are uploaded separately.

Assuming code is submitted:

Project Report and practical software code example	50%
PowerPoint Slides	20%
15 minute YouTube video	15%
2 minute YouTube video	10%
One page summary	5%