Week 1 Journal and In-Class Assignments

Submitted by: Dhiraj Gharana
Submitted to: Dr. Boult (CS 6000)
University of Colorado at Colorado Springs

I. Biography

I'm from Nepal, a country in Asia flanked on either side by India and China. Currently, I am remotely pursuing a Ph.D. in Engineering (concentration in Computer Science) at the University of Colorado at Colorado Springs (UCCS).

Recently, I have been working on a survey paper by the name "A synopsis of cross-situational word learning (CSWL) in robots: with insights from CSWL research in young children," under the supervision of Dr. Jugal Kalita [1]. My computer science research goal is to find better solutions to the problem of robots learning natural languages through multi-modal perception across different contextual situations while conversing and interacting with humans.

My wife, a medical doctor, and I a few months after our marriage. The necklaces she is wearing are the traditional necklaces worn by wives in Nepal. She is also wearing a vermilion red cosmetic powder called "Sindoor," which Hindus believe will bestow longevity and health on her husband. Though I only believe in mathematics and science, I appreciate her display of affection and love.

Apart from research in computer science, I am also interested in contributing to the human endeavor of understanding the mechanisms of strong natural emergence, in which the interaction of a massive collection of microscopic (smaller) entities gives rise to a macroscopic system with new properties and functionalities that were not present in any given single microscopic constituent. Examples of strong natural emergence include the formation of atoms and chemical bonds from the interaction of sub-atomic particles, the creation of multi-cellular organisms, like humans, from the interaction of a myriad of cells, and the existence of an ecosystem from the interaction of many living organisms residing in the biosphere of earth, and more interestingly, the emergence of consciousness due to the interaction of many neurons in the human brain.

A. Goals for This Course

As will be evident from the description provided in section I-C, my initial approach to research cost me significantly in time and effort.

After taking this course by Dr. Boult [2] I hope to learn the canons and the caveats with respect to Computer Science Research so as to develop a cost-effective approach to surveying the literature relevant to my research at hand, to adopt an apt general protocol for conducting, collaborating on, and evaluating research, and to acquaint myself with proper research writing skills and also with the different journals and venues in which I can attempt to publish my work. I am sure that this course will also introduce tools that will prove useful, if not indispensable, to optimally conducting and sharing my research work.

In addition to gaining pertinent knowledge and learning useful skills related to the solution space, as mentioned in the previous paragraph, I believe this course will also guide me in regard to identifying and choosing a proper research problem to pursue and will facilitate my understanding of the various constraints, limitations, and intricacies of the problem space.

B. Research Publications

D. Gharana, S. C. Suh, and M. Kang, "Gender classification based on deep learning" in *Big Data and Visual Analytics*, 1st Ed. S. C. Suh, T. Anthony, Eds. Springer International Publishing, 2017, ch. 3, pp. 55–69, doi: 10.1007/978-3-319-63917-8.

D. Gharana, "Gender and age classification from facial images using deep learning" M.S. thesis, College of Sc. and Eng., Texas A & M University - Commerce, Texas, U.S.A, July. 2016.

C. Academic and Research Lessons Learned

> "*Reading without thinking is useless. Thinking without reading is dangerous.*"
> – Confucius, taken from goodreads, where I replaced 'learning' by 'reading'
> ***https://www.goodreads.com/quotes/9059465-learning-without-thinking-is-useless-thinking-without-learning-is-dangerous***

Any PhD researcher needs to read extensively. Yet, simply reading a lot does not guarantee erudition. Mental retention of facts and acquisition of knowledge are meaningful only if that knowledge can be applied to solve real-life problems. On the other hand, in addition to innovative thinking, successful

[1] Dr. Jugal Kalita is the Chair of the Computer Science Department of the College of Engineering and Applied Sciences at UCCS.

[2] Dr. Boult is El Pomar Endowed Professor of Innovation and Security at UCCS, co-director of the Bachelor of Innovation family of degrees, the chair of the IEEE PAMI Technical Committee, and a founding member of the IEEE Biometrics Council.

research depends heavily on acquaintance to others' work through extensive literature review and collaboration.

The initial part of the above quote is relevant to my undergraduate academic pursuits as I learned from my freshman year that reading for exams and other coursework without an in-depth understanding of the concepts being studied and a vision of how these concepts relate to other previously mastered concepts, never helped me master the subject matter so as to be able to solve real life problems. Consequently, I abandoned this approach early on during my undergraduate studies.

The latter part of the above quote is relevant to my early graduate research endeavors in that zealously pursuing innovative ideas with a dearth of literature review often lead to the duplication of already published work, the cumbersome chore of re-inventing the wheel, the malpractice of coining new terms for technical concepts that already exist in the jargon of published work, and the almost inevitable consequence of having to eventually abandon my research endeavor.

However, I now acknowledge literature review as the most important initial step in understanding the problem at hand and exploring known solutions; the fact that my book chapter research publication is replete with references corroborates this claim.

II. Learning from this week: 09/02 – 09/08

A. My Git Repository

I have a public git repository on GitHub for this course. The following is the url of the repository: **https://github.com/scientiagharana/Dhiraj-Gharana-Dr-Boult-CS6000.git** Dhiraj-Gharana-Dr-Boult-CS6000

1) The Tools Used: I installed git on my system (Windows 10) and then I created the git repository, mentioned in section II-A, online using GitHub. I then cloned the git repository onto my personal system and created the folder "Assignments and Journals," into which I have placed this document. I committed my local changes and then pushed them to the master branch.

B. My Learning

Below, I list the main concepts and skills I learned from this course this week:

1) I learned the qualities a scholarly endeavor must possess in order to qualify as good research.
2) I learned the different approaches used for research in Computer Science.
3) I learned how to use Google Scholar effectively as a research journal search and scan tool.
4) I learned how to evaluate research papers and how to use Google Scholar to view the citation count for any research paper.
5) I was already familiar with Mendeley. Nonetheless, this week I learned about Zotero as a potential alternative for reference management.
6) I learned a bit more about LaTeX and about using LaTeX to write theses and dissertations. However, I have already written research articles in LaTeX using different journal templates and am already adept at using it for simple research writing.
7) This course also taught me the benefits of using Overleaf, especially the Overleaf online LaTeX editor. However, I still prefer writing LaTeX documents offline using TeXstudio's editor, with MiKTeX (the LaTeX typesetting system for Windows), on my own system.

C. Problems and Struggles

1) A research idea becomes research only when shared: In the first video for Lecture 1, Dr. Boult states — research work that is not shared can qualify as merely a research idea and not as research.

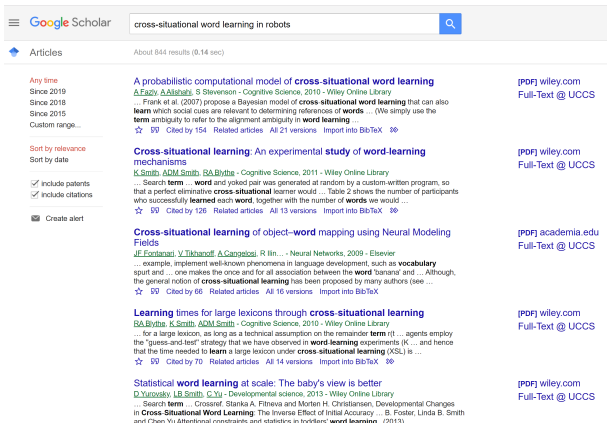
III. Papers Relevant to My Research

IV. In-Class Exercises Related to Lecture 1

A. Google Scholar Exercises

1) Finding A Relevant Research Paper: The research topic I am interested in is "Cross-situational word learning (CSWL) in robots: inspired by CSWL research in young children."

I ran the search key phrase "cross-situational word learning in robots," without quotes, on Google Scholar and obtained the following result.

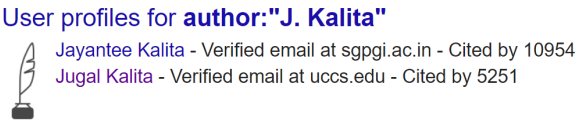


Interestingly, four out of the first five results rendered by Google Scholar, all shown in the figure above excluding "Learning times for large lexicons through cross-situational learning," were very relevant to the survey paper I am writing.

I have used IEEE Xplore, ScienceDirect, and Semantic Scholar to search for journal articles in the past. I am grateful that this course demonstrated the benefits of using Google Scholar for this purpose. I feel I will be using Google Scholar quite frequently from now on.

2) Finding Most Cited Papers by My Supervisor: My research supervisor is Dr. Jugal Kalita.

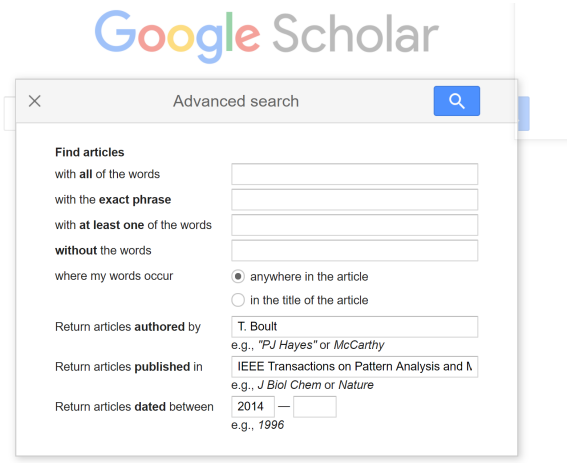
I typed author:"J. Kalita" into the basic search of google scholar, which presented me with two authors to select from, shown below.



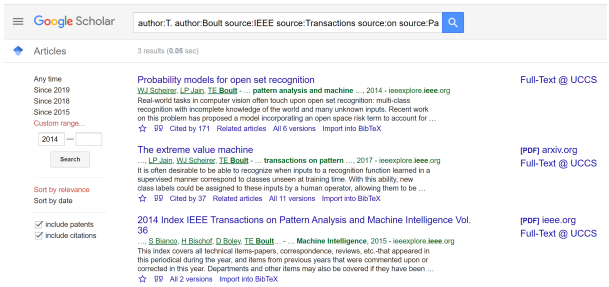
I clicked on my supervisor's name, Jugal Kalita, and reached his profile page, which has his papers sorted by number of citations [3]. The following shows a list of the first few papers listed in Dr. Jugal Kalita's profile page according to their citation ratings.

Title	Cited by	Year
Network anomaly detection: methods, systems and tools [1]	622	2013
Summarizing microblogs automatically [2]	240	2010
Syntactic normalization of twitter messages [3]	195	2010

3) Finding Select Papers Written by Dr. Boult: To find articles authored by **Dr. Terrance Boult**, the instructor of this course, published in *IEEE Transactions on Pattern Analysis and Machine Intelligence* within the past five years, I used the Advanced Search option of Google Scholar as shown below:



The three results, which seems to be the exhaustive list, rendered by Google Scholar is provided below:



I pulled up the citation details [4] for the three papers above, which are provided below.

Title	Cited by	Year
Probability models for open set recognition [4]	173	2014
The extreme value machine [5]	37	2017
2014 Index IEEE Trans. on Pattern Analysis and Machine Intelligence Vol. 36 [6]	N/A	2015

REFERENCES

[1] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: Methods, systems and tools," *IEEE Communications Surveys Tutorials*, vol. 16, pp. 303–336, First 2014.

[2] B. Sharifi, M.-A. Hutton, and J. Kalita, "Summarizing microblogs automatically," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, HLT '10, (Stroudsburg, PA, USA), pp. 685–688, Association for Computational Linguistics, 2010.

[3] M. Kaufmann and J. Kalita, "Syntactic normalization of twitter messages," in *International conference on natural language processing*, (Kharagpur, India), 01 2010.

[4] W. J. Scheirer, L. P. Jain, and T. E. Boult, "Probability models for open set recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 11, pp. 2317–2324, 2014.

[5] E. M. Rudd, L. P. Jain, W. J. Scheirer, and T. E. Boult, "The extreme value machine," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 3, pp. 762–768, 2017.

[6] R. Abiantun, N. Ahuja, A. Aissani, K. Ait-Mohand, K. Aizawa, Z. Akata, E. Akbas, A. Alahi, J. Almazan, L. Alvarez, *et al.*, "2014 index ieee transactions on pattern analysis and machine intelligence vol. 36," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 1, p. 209, 2015.

[3] I assume that the greater the number of citations a paper receives, the better the quality of research presented in that paper

[4] "2014 Index IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 36" is an index into the authors who presented and the subjects that were presented in this particular conference. Hence, citation count of this index paper is irrelevant, for which the symbol N/A has been applied.