

CPSC 490 Proposal Fall 2015

## Lien Tran Class of 2016

*Professor Dana Angluin - Department of Computer Science*

*Professor Robert Frank - Department of Linguistics*

### OVERVIEW

This project aims to address the problem of authorship verification through the use of deep learning. Through this project, the student will explore two specific topics: authorship attribution in linguistics, and machine learning in computer science and statistics. To apply this knowledge, the student will implement a learning model to analyze corpuses of text from famous literary authors, and learn to assess whether other pieces of writing were written by the chosen author. The implementation results will be compared with those of previous attempts to do so in existing literature in linguistics.

### BACKGROUND

Authorship attribution is the problem of analyzing and characterizing an individual's writing style in order to determine the most likely author for an unlabeled body of natural language text. In the past century, both linguists and statisticians alike have addressed the problem amply. Methods of analysis span from linguistic approaches, such as syntactic and semantic analysis, to statistical and computational techniques employing probability models and support vector machines. Among these, deep learning has recently attracted significant interest, as well as showed promising results.

Deep learning is a subset of statistical methods that fall under the field of machine learning. It offers the ability to capture high-level abstractions through unsupervised or semi-supervised feature extraction. In the past decade, numerous applications involving deep learning have born impressive results in various fields, including computer vision, bioinformatics, and other data-driven industries. In the field of linguistics, and particularly in applications for authorship attribution, researchers have only begun to apply deep learning in the past couple of years.

### PROJECT DESCRIPTION

Through the course of a semester, the project will consist of primarily two parts: research deep learning models, and implement a model in order to address a specific type of author attribution task.

The first half of the semester will be dedicated to researching current approaches to solving the author attribution problem and exploring deep learning methods. Several relevant papers on authorship attribution have already been collected<sup>1-5</sup>, and will be examined carefully in the first few weeks of the semester. As for deep learning, the student will use online lectures, tutorials and other academic resources in order to establish a workable understanding of this highly technical material.

The second part of this project will involve building a learning model to solve a particular problem in authorship attribution. More specifically, this model should learn to recognize

patterns from a dataset of literary writing from one author of choice, ie. Ernest Hemmingway, and then, given a vanilla piece of natural language input, output whether that input text belongs to the learned author's writing. This task is called authorship verification: the learned model verifies whether a piece of writing belongs to a target author.

Part of devising the learning model will also include collecting a workable dataset and coming up with a system for evaluating results. For these tasks, Koppel and Schler's 2004 paper on *Authorship Verification as a One-Class Classification Problem*<sup>5</sup> will be used, as it describes a highly accurate method to solve this binary classification problem using support vector machines. The project's final implementation of a learning model will use datasets including, but not limited to, the one mentioned in Koppel and Schler's study. Final results from the model will be evaluated against Koppel and Schler's accuracy, as well as other possible sources that may arise during research.

## OBJECTIVES

- ❖ Explore deep learning through online courses, tutorials and literature in statistics and computer science
- ❖ Learn and examine authorship attribution methods for literary works
- ❖ Research and collect a suitable dataset to facilitate results evaluation
- ❖ Implement, either with or without assistance of external software packages, a deep learning networks to perform authorship verification for literary works in the English language
- ❖ Compare results to similar attempts on authorship attribution in the literature

## EDUCATIONAL MATERIAL

UFLDL Tutorial: [http://deeplearning.stanford.edu/wiki/index.php/UFLDL\\_Tutorial](http://deeplearning.stanford.edu/wiki/index.php/UFLDL_Tutorial)

Torch Scientific Computing Framework: <http://torch.ch/>

Caffe Deep Learning Framework: <http://caffe.berkeleyvision.org/>

## REFERENCES

1. Stamatatos, E. 2009. A Survey of Modern Authorship Attribution Methods.
2. Koppel, M., Schler, J. and Argamon, S. 2009. Computational Methods in Authorship Attribution
3. Stanczyk, U. and Cyran, K. 2007. Machine Learning Approach to Authorship Attribution of Literary Texts.
4. Diederich, J. 2003. Authorship Attribution with Support Vector Machines.
5. Koppel, M. and Schler, J. 2004. Authorship Verification as a One-Class Classification Problem.