Annie En-Shiun Lee
| Teaching Dossier

Table of Contents

Statement of Teaching Philosophy

Introduction

In this age of the knowledge economy, Canada is transforming itself to become a global leader in research innovation and excellence. Given this outlook, the central theme of my teaching philosophy focuses on training the next generation of exceptional minds. My teaching philosophy is divided into three themes of Atmosphere, Application, and Attention (A-cubed)

Inclusive and Supportive Atmosphere:

The heart of my administrative service lies in creating an inclusive and supportive environment for training incoming researchers. The organization of the two workshops, Women in Machine Learning (co-hosted with NeurIPs 2010) and Broadening Participation in Data Mining (co-hosted with SIGKDD 2017), utilizes my natural talents in coordination. I was able to strategically create a volunteer system as well as orchestrate the mentoring program. Furthermore, throughout my career I have diligently yet emphatically advised and mentored students of all levels (high school to masters) and consulted team projects (course projects, undergraduate design projects, and continuing education capstone project). When studying interpersonal dynamics, Google found that psychological safety is a key component of building a strong relationship[1]. Likewise, my collaborator's masters student acknowledged that I "played a decisive role in shaping their view on research and the importance of being precise and diligent" and provided "encouragement and patience throughout the duration of their research".

Real-World Application:

My strongest demonstration of problem-based learning is from my course, Applied Machine Learning and Lifecycle, in the award-winning Machine Learning Certificate program at York University. Overall, the course project is broken down into milestones that are supported by materials from the weekly learning activities. Within the first two in-person classes, the learners are motivated with the real-world problem of text classification and provided with the necessary code tutorials to jump-start their projects. In order to keep projects on-track, I dedicated office hours, milestone updates, and flexible one-on-one consultations. These activities are crucial to teach learners real-world skill sets, such as organizing their projects into deliverable, seeking feedback, and communicating to stakeholders. When Carnegie Mellon Professor, Randy Psauch, had his students design a virtual world, he was blown away by their results[2]. He "didn't know how high the bar should be, and he'd only do (the students) a disservice by putting it anywhere." Similarly I was blown away by the project accomplishments. The learners report a high satisfaction rate despite the course's difficulty and the large amount of materials covered.

Captivate Attention:

In today's digital age, capturing the learner's attention is crucial. When classes transitioned online during COVID-19, I rapidly pivoted into online teaching in an agile fashion. As an early-adopter, I designed, tested out, and incorporated synchronous techniques of engaging online learners (such as icebreakers, polling, breakout rooms, and movement breaks). I experimented these techniques with learners of all ages, ranging from lower and middle elementary grades to adult learners. Further demonstrating innovative leadership, I shared these advance technique with instructional teams of various programs, thus encouraging others to adopt them into their online instruction. Another support of my scholarly innovation in teaching is my pedagogy paper on context-aware feedback[3]. This paper incorporated context-aware feedback to explore the variables that maximize student responses to smartphone requests in various learning contexts. The results demonstrate that flexible conditions from the out-of-class condition ranked higher on the active theory scales than the rigid in-class condition.

Duhigg, Charles. "What Google learned from its quest to build the perfect team." The New York Times Magazine 26 (2016): 2016.

Pausch, Randy. The last lecture. Hachette Books, 2008.

Lee, En-Shiun Annie. "Applying Activity Theory of Mobile Learning to Context-Aware Smartphones." E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Association for the Advancement of Computing in Education (AACE), 2012.

Teaching Experiences

Course Developer and Instructor at York University

CSML 1010: Applied Machine Learning and Lifecycle — **Certificate in Machine Learning**
Course Developer and Instructor, 20 students — *Fall2018, Winter2019, Fall2019, Winter2020*
design course content for 8 week curriculum; write weekly update emails; teach three in-person classes; host weekly office hour and project consultation; mark project milestones and code exercises.

Teaching, Instructional, or Lab Assistant at University of Waterloo

MTE 140: Algorithms and Data Structures — **Mechatronics Engineering**
Lab, Tutorial, and Evaluation Teaching Assistant, 98 students — *Spring 2010, 2011*
co-author of course assignment and textbook; teach weekly labs and host office hours; host midterm and final review sessions; create scripts to mark course projects; mark midterm and exam marking

SYDE 422: Machine Intelligence — **Systems Design Engineering**
Review and Evaluation Teaching Assistant, 33 Students — *Winter 2011, 2012, 2013*
teach biweekly tutorials; give guest lecture on evolutionary algorithm; host project consultation office hours; use LEARN to group projects, create and grade quizzes, and enter marks.

SYDE 223: Data Structures and Algorithms — **Systems Design Engineering**
Review and Tutorial Teaching Assistant, about 80 students — *Winter 2010*
teach weekly tutorials; mark exams; host office hours.

SYDE 114: Introduction to Calculus — **Systems Design Engineering**
Tutorial and Evaluation Teaching Assistant — *Fall 2012*
teach bi-weekly tutorials; host first-year office hours; mark assignments and tests; enter marks using LEARN platform.

SYDE 113: Linear System and Matrices — **Systems Design Engineering**
Tutorial and Evaluation Teaching Assistant, 113 Students — *Fall 2010, Fall 2011*
teach bi-weekly tutorials; host first-year office hours; mark quizzes and tests.

CS134: Principles of Computer Science — **School of Computer Science**
Tutorial and Evaluation Instructional Assistant — *Fall 2007, Winter 2008*
teach weekly tutorial sessions in two sessions; attend weekly course meetings; mark assignments, midterms, and exams.

CS133: Developing Programming Principles — **School of Computer Science**
Lab and Evaluation Teaching Assistant, 20 to 30 Students — *Fall 2006*
attend weekly labs; mark assignments, midterms, and exams.

CS240: Data Structures and Data Management — **School of Computer Science**
Graduate Teaching Assistant — *Spring 2007*

CS230: Introduction to Computers and Computer Systems — **School of Computer Science**
Graduate Teaching Assistant — *Winter 2007*

Undergraduate Marker at University of Waterloo

MATH235: Linear Algebra 2 — **Department of Mathematics**
Undergraduate Marker — *Winter 2003, Fall 2003, Winter 2004*

MATH239: Introduction to Combinatorics — **Department of Mathematics**
Undergraduate Marker — *Fall 2001*

Other Experiences

Reviewer/Program Committee: ACM SIG Computer Science Education 2013-2021, Women in Machine Learning Workshop Unworkshop co-hosted with ICML 2020, Workshop on Broadening Participation of Data Mining co-hosted with SIGKDD 2017, Women in Machine Learning Workshop co-hosted with NeurIPS 2010/2018

Supervisor: CSML1030 Machine Learning Capstone 2020 - Pattern Analysis on Biosequence, Fields Summer Undergraduate Research Program 2018 – Big Data Extraction, 2nd and 4th year design projects, undergraduate research assistant (5 times)

Instructor: Learn Easy 2020 - Artificial Intelligence and Machine Learning for Kids/Tweens/Teens, National Learning Code 2017-2019 - Introduction to Artificial Intelligence and Machine Learning, Ladies Learning Code: Data Insights with Python

Speaker/Panellist: Windsor-Essex DevFest 2019, AI with the Best 2016 & 2018, Sentiment Symposium 2017, Canadian Applied and Industrial Mathematics Society Annual Meeting 2017, Toronto Machine Learning Summit 2017, Fields Industry Lunch Seminar 2017, Conference on Big Data and Information Analytics 2017, Ladies who Tech - Filling the Gap 2017

Language Proficiency: English (Native Proficiency), Mandarin (Professional Working Proficiency)

Evidence of Teaching Effectiveness

List of Awards

- 2020 Canadian Association for University Continuing Education Program Award - Non-credit Programming over 48 hours
- Nomination for Amit and Meena Chakma Award for Exceptional Teaching by a Student, University of Waterloo, 2011

Feedback from Students

Supervision of Masters Research

- The completion of this study would not have been possible without the support and nurturing of Dr. Annie Lee. You played a **decisive role** in shaping my view on research and the **importance of being precise and diligent**. Thanks for providing me with **encouragement and patience** throughout the duration of my research. – Afsaneh Towhidi, August 2019, Masters of Applied Science in Electrical and Computer Engineering, University of Ontario Institute of Technology
- I would like to thank Dr. Annie En-Shiun Lee, for her **tremendous leadership and support** when I was her research assistant. Without her support, I would not have **decided to partake in my Master's journey**. Even during my Master's, she was a huge support for me. – Sanderz Fung, 2015, Master of Applied Science in Systems Design Engineering

Teaching CSML1010 Applied Machine Learning Life-Cycle

- The **remote Zoom meeting format is great** and would recommend the course to colleagues and friends. – Student, Winter 2020
- This course was **rich with course material, reading links and example codes**. It was **nicely organized with clear instructions**. – Student, Fall 2019
- I **learned a lot doing the project and the coding exercises** helped me **learn how to continue**. – Student, Fall 2019

Teaching MTE140 Algorithms and Data Structures

- Annie has in depth working knowledge of required concepts and is able to **filter out required information** and give students a **succinct explanation** of a particular concept. – Student, Spring 2010
- Annie developed a good rapport with the students and when she was explaining something to them, she **explained it well, and at a level they could understand**, which is very important. – Instructional Staff, Spring 2010
- Annie is a very **confident speaker** who is capable of doing comprehensive presentations with little preparation. – Student,Spring 2010

Feedback from Colleagues

Organization and Coordination

- Dr. Lee and I were students of the same Ph.D. adviser. I am particular impressed by Dr. Lee's **ability to manage multiple research projects concurrently**, **each project resulting in a separate publication with novel scientific contributions**. Dr. Lee **pays attention to the team members' needs**, thus building a strong team. The thing that stood out with Dr. Lee is her **superbly organization and outstanding presentations**; any employer will be lucky to have her as a research scientist. – Gary Li, Ph.D. October 18, 2017
- I worked with Annie on many large and small projects, and enjoyed every minute of every one of them! Annie always **brought a great deal of knowledge, enthusiasm and positive energy into a project**, **with much patience in explaining details and answering questions**. On all occasions, she took the feedback very seriously, and used it to improve her already great work. Annie's **exceptional liveliness and generosity** always put a smile on everyoneâĂŹs face, especially during our Monday morning team meetings. I will especially miss our early morning brain-storming sessions over coffee. Annie is a true scientist with brilliant ideas, a real team player, a wonderful friend, and will be a great addition to any research team. –Afsaneh Fazly, May 16, 2017
- Annie possesses the acumen to **effectively strategize and plan research projects of any scale and complexity**! I had the pleasure of working with her for more than a year at VerticalScope, where she **led the RD team's research efforts** on Machine Learning, Natural Language Processing, and Social Network Analysis. I closely worked with her on projects involving Named Entity Recognition and Linking, Content Recommender, and User Influence. Throughout all these projects, she was **meticulous and organized in her research methods**, where she carefully filtered, analyzed and summarized state-of-the-art research advancements. Annie also came up with **creative and novel research ideas for the team to quickly test concepts and implement prototypes**. Moreover, she has **the knack for laying out well-planned project roadmaps which were pivotal for the engineering team's success** in all these projects. Annie was **never afraid to take initiative**; I was especially impressed with her proactiveness in the Topic Modelling project wherein she reached out to domain experts outside of our team in order to get valuable assistance and feedback on our work. Annie has a unique mix of skills involving people, processes and technology that will make her a great asset for any company. – Krish Perumal, September 26, 2018

Public Speaking

- **Great public speaking /lecturing** skills (Professor, ExpecTAtion)
- Love the **enthusiasm and personal stories**, like that (the summary is) related to the University of Waterloo (Colleague, ExpecTAtion)

Evidence of Professional Development

Relevant Publications

1. **En-Shiun Annie Lee**, Karthik Kuber, Hashmat Rohian, and Sean Woodhead. Creating a problem-based machine learning curriculum for the flipped online setting. In *In Progress*.

2. **En-Shiun Annie Lee**. Applying activity theory of mobile learning to context-aware smartphones. In *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, volume 2012, pages 1092–1098, 2012.

Books, Book Chapters, and Reviews

1. **En-Shiun Annie Lee**, Ho-Yin Sze-To, Andrew K. C. Wong, and Daniel Stashuk. Unsupervised pattern discovery in biosequences using aligned pattern clustering. *SM Journal of Bioinformatics and Proteomics*, 1(2):1008, 2016.

2. Alexander Wong, James R. Wallace, **En-Shiun Annie Lee**, Xiao Yu Wang, Victor Cheung, and Abhishek Kumar. *Data Structures and Algorithms in a Nutshell*. University of Waterloo, 2012.

3. Andrew K. C. Wong, Dennis Zhuang, Gary CL Li, and **En-Shiun Annie Lee**. Pattern discovery and recognition in sequences. In *Pattern Recognition, Machine Intelligence and Biometrics*, pages 29–59. Springer, 2011.

Preparation of Teaching - Certificate for University Teaching

The Certificate in University Teaching (CUT) program is offered by the Centre for Teaching Excellence at the University of Waterloo. There are three courses required for the program, each is described below.

Certificate for University Teaching GS901 Preparing for University Teaching:

In Graduate Studies (GS) 901, I built reflection skills, theoretical knowledge, and applications by attending six workshops about teaching and learning in higher education and write reflective, application-based response papers. Each of the workshop is about 1.5 hour in length and all of them are listed below with the title of the corresponding response paper. The only exception is the two full day ExpecTAions workshop offered by the Faculty of Engineering. I have subsequently functioned as graduate trainer in the workshop and offered advice and practical experience to new teaching assistants. The 6 workshops I took in GS901 are as follows:

1: Course Design: Detailed Alignment Chart

2: Understanding the Learner: Personal Unique Learning Style

3: Designing Exam: Exam Questions for Future Teaching

4: Critical Thinking: Application, Assessment, and Collaboration

5: Building Credibility in a Teaching Environment

6: Faculty of Engineering ExpecTAtions Workshop Training

Certificate for University Teaching GS902 Preparing for University Teaching:

In Graduate Studies (GS) 902, I investigated higher education theory, further developed applications to communicate to a general university audience, and articulated my self-awareness as a teacher. I complete two projects: a research project about a teaching and learning issue in higher education, and

a teaching dossier. The research project, "The Use of Smart Phone Technology to Facilitate Deep Learning", was a 10-page report presented at a teaching conference and published as a conference paper.

Certificate for University Teaching GS903 Teaching Practicum:

In Graduate Studies (GS) 903, I practiced and received feedback on my individual teaching skills and become more self-aware and critically reflective. I had three of my teaching observed and received feedback from a trained observer, and write a reflective response paper on the feedback received for each observation. Below are the three listed review sessions and comments from the trained observer.

MTE 140: Algorithms and Data Structures **Mechatronics Engineering**
Lab, Tutorial, and Evaluation TA, 98 students *Midterm Review Session*

- Presented extemporaneously as a confident, knowledgeable, and friendly instructor
- Effectively organized lecture by alternating between PowerPoint presentation and blackboard

SYDE 422: Machine Intelligence **Systems Design Engineering**
Review and Evaluation TA, 33 Students *Tutorial Session*

- Built rapport with students through personable introduction and enthusiastic, relatable description of research
- Use of classroom mechanics such as the one-minute summary and interactive group work

SYDE 113: Linear System and Matrices **Systems Design Engineering**
Tutorial and Evaluation TA, 106 Students *Introductory Tutorial Session*

- Projected confidence and established credibility in a large class using enthusiastic body language that is dynamic and exciting

Sample Teaching Materials

Course Syllabus for CSML1010 Applied Machine Learning and Lifecycle

COURSE OUTLINE | SESSIONS | DISCUSSIONS | ASSESSMENTS | HELP

Hello, and welcome to Applied Machine Learning and Life Cycle, the second course in York University's blended Certificate in Machine Learning!

My name is Annie Lee, and I will be your instructor.

I hold a PhD from the University of Waterloo at the Centre of Pattern Analysis and Machine Intelligence with more than 12 years of experience in data mining and machine learning. My passion for finding patterns in society and in nature in the big data era has lead to dozens of publications in computational advertising, sentiment analysis, and sequence analysis, including a highly-read review paper on big data. Most notably, I developed explainable unsupervised algorithms, which uncovered patterns utilising clustering and partitioning of raw data with *a priori* knowledge.

Currently, I am a Research Scientist working on Natural Language Understanding at Stradigi AI. During my tenure, I have worked on novel algorithms such as deep neural pruning for language models and adversarial transfer learning. More specifically, I specialize in sequential tagging neural architectures, applied to Named Entity Recognition and Key Phrase Extraction.

In the past, I have served in Broadening Participation in Data Mining, co-hosted with Knowledge Discovery and Data Mining (SIGKDD), and Women in Machine Learning, co-hosted with Neural Information Processing Systems (NeurIPs).

Contact Information: ealee@uwaterloo.ca.

Print CSML1010 Course Outline

COURSE DESCRIPTION

This course covers the iterative life-cycle of machine learning, including processing data, data exploration, feature engineering, feature selection, model selection, parameter optimization, performance evaluation, and model interpretability. We will tackle the unstructured data of text and cover advanced techniques for improving its performance. You'll gain hands-on experience using Python to solve text classification and construct advanced algorithms such as feature selection and ensemble methods.

MACHINE LEARNING

OVERALL LEARNING OUTCOMES

Upon successful completion of this course, learners will have acquired the following knowledge, skills, or attitudes to:

1. Plan each step of the machine learning life-cycle by identifying the practical problems, determining the metrics that define a successful outcome, evaluating constraints and shortcomings, deciding on appropriate approaches by selecting several potential machine learning models, and building a minimal viable product prototype
2. Explore data (structured and unstructured) by performing basic exploration such as visualization and analyzing the data, computing exploratory statistics, extracting and design features, and carry out further feature engineering techniques such as feature selection or feature creation.
3. Select the appropriate machine learning models to match the type of problem, know the trade-off (shortfalls and strengths) of each model, compute proper evaluation metrics and learning curve, interpret the model and explain the results, iterate on models strategically to improve performance (i.e. from the baseline model to the ensemble model).

COURSE FORMAT AND METHODS OF LEARNING

This course has been designed to balance acquiring knowledge with applying them in a practical, meaningful way. This will be done by encouraging class participation and interactivity. Various teaching techniques will be used in this course to facilitate learning, including lectures, group discussions, assignments, and on-going problem solutions and case studies.

MARKING BREAKDOWN

Note that students MUST pass the individual components to pass the entire course.

ASSESSMENT	% OF FINAL GRADE
Project (Individual or Pair) - Proposal, Milestone 1, Milestone 2, Final Project Submission (Presentation)	60 (10, 10, 10, 30)
In-Class Code Exercise (for the first two days in Jupyter notebooks)	10 (5 each)
Weekly Online Update in Forum Discussion - in-depth, quality, insightful	10
Weekly Online Code Exercises (Jupyter notebooks)	10
Independent Learning Presentation	10

Sessions

COURSE OUTLINE | SESSIONS | DISCUSSIONS | ASSESSMENTS | HELP

For details on any particular session, click on the session links below.

CLASS NUMBER AND DATE	LECTURE READINGS AND TOPICS	ASSIGNMENTS AND ACTIVITIES
IN-PERSON SESSION 1 Saturday, March 21, 2020 Time: 10:00 AM - 5:00 PM Location: LSB101 (Life Sciences Building Room 101)	INTRODUCTION TO THE MACHINE LEARNING CYCLE • Introduction to Machine Learning Lifecycle • Data Exploration and Visualization • Feature Extraction and Feature Selection • Combining Features and Reducing Dimensions	Introduction to the Machine Learning lifecycle and discussions on sample business problems. Students will set-up their programming environment and have time to work on their own individual projects. In-class code exercises to be handed in the end of the day.
IN-PERSON SESSION 2 Sunday, March 22, 2020 Time: 10:00 AM - 5:00 PM Location: LSB101 (Life Sciences Building Room 101)	INTRODUCTION TO THE MACHINE LEARNING CYCLE (CONTINUED) • Models Evaluation and Assessment • Iterative Improvement • Ensemble Algorithms • Researching methods and papers	Further steps of the Machine Learning lifecycle, especially focusing on the models. Students will have time to work together with three different people. In-class code exercises to be handed in the end of the day.
ONLINE SESSION 1 March 23 - 29, 2020 Due Sunday, March 29, 2020	Introduction to Text Classification	Weekly Progress Report Coding Exercise Project Proposal and Exploratory Data Analysis Due
ONLINE SESSION 2 March 30 - April 5, 2020 Due Sunday, April 5, 2020	Feature Extraction	Weekly Progress Report Coding Exercise(imbalanced data) PResentation 1: Project Proposal
ONLINE SESSION 3 April 6 - 12, 2020 Due Monday, April 13, 2020	Feature Selection	Weekly Progress Report Coding Exercise Project Milestone 1 Due (submit then One-on-One)
ONLINE SESSION 4 April 13 - 19, 2020 Due Monday, April 20, 2020	Model Selection by Evaluation Metrics and Learning Curves	Weekly Progress Report Coding Exercise Presentation 2: Present Project Milestone 1
ONLINE SESSION 5 April 20 - 26, 2020 Due Sunday, April 26, 2020	Ensemble Learning Methods	Weekly Progress Report Coding Exercise Project Milestone 2 Due (submit then One-on-One)
ONLINE SESSION 6 April 27 - May 3, 2020 Due Sunday, May 3, 2020	Interpreting and Explaining Models	Weekly Progress Report Coding Exercise Presentation 3: Present Project Milestone 2
Optional Content May 4 - 21, 2020	[OPTIONAL] Deep and Sequential Algorithms [OPTIONAL] Active Learning and Transfer Learning [OPTIONAL] Reinforcement Learning	[Optional] Preliminary Project Feedback
Friday, May 22, 2020 Time: 10:00 AM - 5:00 PM Location: SC302	Project Final Presentations / Course Wrap-up	Project Final Due Project presentations (submit and class presentation) Independent Learning Presentation

Last modified: Wednesday, 1 April 2020, 4:51 AM

ASSIGNMENT DETAILS

PROJECT PROPOSAL: Problem, Dataset, Exploratory Data Analysis

1. Find a Text Classification problem in the area of natural language processing or information retrieval.

2. Problem Selection and Problem Definition: Select a Text Classification problem from natural language processing or information retrieval problem. Define the problem (input and output), decide on several basic methodologies and identify the limitations, shortcomings, constraints of each method

- Natural Language Processing
- NLP-progress
- SemEval2018, SemEval2017, SemEval2016, SemEval2015, SemEval2014, SemEval2013, SemEval2012

3. Dataset: Find a datasets for your problem. Only one project per dataset (first come first serve). Special consideration (marks) will be given to more difficult problem definition and non-standard datasets.

- Open Datasetlist and look under NLP
- Open Datasets, look under "Text Datasets"
- Additional Dataset from University of Melburn
- Google Dataset Search

Additional Datasets by NLP Problems:

Automatic KeyPhrase Extraction: Raw Data, Process Data
Named Entity Recognition: WNUT (noisy text), CONLL (news), Finance
Topic Classification: topic labels, news, newsgroups
Hate Speech: Twitter
Fake News: news articles
Text Corpus: Enron Dataset; Google Book ngrams; Blogger Corpus; Wikipedia Links Data; Gutenberg E-books; Hansard Canadian Parliament; "Financial Statement Data Sets"
Question and Answering: Jeopardy Questions;
Spam: SMS Spam; UCI Spam Repository;
Sentiment Analysis: Multi-Domain Sentiment; IMDB-sentiment; Stanford Sentiment; Sentiment 140; Twitter Airline Sentiment; Sentiment Analysis in Twitter; Fine-Grained Sentiment Analysis on Financial Microblogs and News;
Review Dataset: Yelp dataset; Amazon Reviews;

4. Data Preparation and Data Exploration: clean, prepare and explore your data.

Reduce your dataset size: For those projects that have datasets that are 500,000+ rows/records, that is too BIG. Aim to have around 2,000 to 5,000 for your first round of iteration. Also, remember that to save time, you can save your pre-processed cleaned data/features (using pickle or output files csv/json/xml) and model checkpoints. It is time-consuming to work on such a large model right away on your first iteration (especially without cloud computing or GPU); that can be saved for your next course. Learning Google Cloud Platform is not an objective of this course. Note that GCP or deep learning will be covered fully in great detail in your next course. ML1020 will assume that learners has no deep learning or cloud background, and thus will have detailed instructions on setting up services (clusters, etc) on GCP.

Balance out your dataset: Most projects are working with unbalanced labels (i.e. 95% in one class and 5% in another class). Please make sure to balance out your class labels, there are examples you can follow in Online Session 4 (extra jupyter notebook exercises). Most datasets have unbalanced labels (i.e. 95% in one class and 5% in another class). Please make sure to balance out your class labels, there are examples you can follow in Online Session 4 (extra jupyter notebook exercises).

5. Submit the code the above steps and create a PowerPoint presentation summarizing your work. Book a 10 minute presentation time during office hours with the instructor to go over your project.

Your project proposal will be marked on the following scheme:

1. Problem Definition (/5)
2. Data Cleaning and balancing dataset (/5)
3. Data Exploration (/5)
4. Others (presentation, graphs/charts/tables, code) (/5)

PROJECT MILESTONE 1: Feature Engineering and Feature Selection

1.Feature Engineering: Perform feature extraction AND feature selection on your dataset. More specifically, 1) perform Feature Extraction by trying different types of features; AND 2) perform Feature Selection by selecting the best features. Marks will be allocated for each. Remember to select ONE specific baseline machine learning model (i.e. such as decision tree, or logistic regression) to benchmark the performance of these different features and feature selection techniques

ASIDE: For advanced feature engineering, consider using "word embedding" (word2vec, Glove), as well as newer embeddings such as "language models" like FLAIR, BERT, ELMO. You may find a quick tutorial on how to use them HERE .

2. Submit a PDF of your jupyter notebook and prepare a presentation of what you have accomplished. Book a presentation time during office hour with the instructor to go over your project.

MARKING RUBRICS

Here is how different elements of the course will be assessed.

Rubric for Weekly Online Update

	Poor Below Standards	Fair Approaching Standards	Good Meet Standards	Excellent Exceed Standards
Contributing to Online Discussions	Attended, but was mostly in listening mode. 0 to 5 points	Attended and responded to questions only. 6 points	Attended, asked questions of class members and instructor, and responded to others. 7 to 8 points	Attended, was fully engaged, asked questions of class members and instructor, responded to others, and generally contributed to the learning of the group. 9 to 10 points

Rubic for In-Class Code Exercise and Weekly Online Code Exercises

Assessment Criteria	Not Good Enough (Score between 0 and 3)	Good (Score between 3 and 4)	Very Good (Score between 4 and 5)
Interpretation of Data (qualitative)	Little or no attempt to interpret data; or there are significant errors; or some data are over- or under-interpreted.	Interpret most data correctly; part of conclusions may be suspect; suggestions on future implementation are sound.	Data are completely and appropriately interpreted; there is no over- or under-interpretation; draw convincing conclusions.
Analysis (quantitative)	Methods are completely misapplied or applied but with significant errors or omissions. Choose inappropriate methods and make wrong predictions.	Most statistical methods are correctly applied but more could have been done with the data. Predictions are sensible but may deviate from the true results in a large range.	Statistical methods are fully and correctly applied; demonstrate superior data analysis skills; deeply mine the data and obtain useful insights for decision making.
Critical evaluation of findings	Blindly accept defective results; or recognize defective results but does not know how to fix them.	Recognize defective results and figure out the causes; understand the main sources of errors.	Show deep understanding for the sources of errors; recognize defective results and eliminates the causes.
Ability to draw proper conclusions and make effective suggestions	No drawn conclusions; draw incorrect conclusions; suggestions are not acceptable.	Draw correct conclusion; suggestions may have potential impact on the future business.	Demonstrate substantial understanding of the problem; conduct deep data analytics using correct methods; draw correct conclusions with sufficient explanation and elaboration.

Rubric for Course Project - Proposal, Milestone 1, Milestone 2, Final

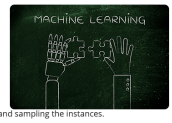
Assessment Criteria	Poor (Score between 0 and 5)	Fair (Score between 6 and 7)	Good (Score between 7 and 8)	Excellent (Score between 9 and 10)
Deep understanding of theory and its applications using qualitative methods to answer problem posed	Demonstrate inadequate understanding of important concepts, methods or their applications, e.g., choose wrong methods, conduct analysis inappropriately, or interpret results incorrectly.	Understand concepts and methods relatively well, analyze data using acceptable methods although not perfect; be able to derive useful information for decision making.	Understand concepts and methods well, analyze data using acceptable methods; know the exact scopes and possible limitations of each method; show capability of using data analytics skills to make right methodology decision.	Demonstrate sophisticated understanding for the concepts and methods; incorporate a state-of-the-art recent method.
Implementation and interpretation of data analysis techniques	Use wrong techniques to analyze data, present inappropriate interpretations or conclusions.	Choose acceptable methods to analyze data, interpretations are sensible, derive useful results.	Use appropriate techniques to analyze data, interpret the results correctly, draw right conclusions based on data analysis.	Use advanced techniques to conduct thorough and insightful analysis, interpret results are in-depth and novel.

Online Session 1 Moodle Material for CSML1010 Applied Machine Learning and Lifecycle.....

Online Session 1: Introduction to Text Classification

COURSE OUTLINE | SESSIONS | DISCUSSIONS | ASSESSMENTS | HELP | 🏠

OVERVIEW



Text classification is a common method of machine learning that is applied to text. It can be found in a large variety of applications, such as sentiment analysis, spam or fraud detection, topic (i.e. subject, genre, category) classification, as well as identifying bullying, fake news, derogatory posts, authorship, age, and gender. As you can see, text classification has wide varying applications that cover any domain that contains text, including social media, financial documents, websites, news articles, and many more.

This week's module focuses on an overall introduction to Text Classification, specifically focusing on the steps of the machine learning life cycle. Your tasks will include 1) execute the coding exercise/assignment and 2) look for dataset for your project (which is worth 60% of your final grade), and 3) answering the 3 weekly questions in the discussion section. First, you will improve your Python skills by executing the jupyter notebooks from github repositories as well as install the different tools and libraries. You will learn how to perform pre-processing and exploratory data analysis on text data and apply end-to-end text classification techniques. For the project proposal, you will select a dataset and begin cleaning the raw data and sampling the instances.

Upcoming this week...

- Remember to submit your in-class code exercises from In-class Day 1 and Day 2 (5% each, total 10% of your final grade).
- We only want a couple of people working on sentiment analysis and no duplicate project on the same datasets. Please put a "reservation" on your project problem and dataset (first come first serve) by replying to the thread on Moodle Discussion, under "Moodle Discussion Forum"-->"General Course Questions"--> "Your problem and dataset (first come first serve)". Also let us know whether the project is individual or pair (whom your partner is).
- Please book a one-on-one meeting with the instructor to get approval for your project. To get my approval and go-ahead for your Problem Definition and Dataset, please book a meeting with me this week ASAP (before your first Project Proposal deadline). Please book a meeting with me to finalize your problem and dataset.
- Also remember to setup your project group submission folders. This needs to be set-up next before your project proposal deadline so that you can submit your work.
- The individual project proposal is due, please read the course main page for what is expected and mark breakdown.

It was great to meet everyone this weekend and get started on our journey together. Hopefully, everyone is now up and running on Python and will be well on their way for the code exercises and project. Best of luck with your learning this week. Any questions along the way, please reach out and ask via the General Course Questions discussion forum.

LEARNING OUTCOMES

By the end of this module, you will be able to describe the following about data mining:

1. Explain the problem and solution of text classification.
2. Give applied examples of text classification.
3. Describe and code a text classification application in Python in order to practice solving the problem.
4. Apply data exploration of text classification on their own project problem in order to gain practical experience.

READ

(The following links will open in new window)

1. Multi-Class Text Classification with Scikit-Learn. Susan Li Feb 19, 2018 · 11 min read

This article and Jupyter notebook solves a multi-class text classification problem with scikit-learn. It follows one iteration of the machine learning life cycle by: 1) formulating the problem, 2) exploring the data (include handling Imbalanced classes), 4) feature engineering (by representing the text and designing feature), 5) model selection and evaluation. Hopefully, it will give an overview of one iteration of the project without focusing too much on the different types of models.

2. Modern Text Mining with Python, Part 1 of 5: Introduction, cleaning and linguistics. datanizing GmbH Mar 24 · 7 min read

Modern Text Mining with Python, Part 2 of 5: Data Exploration with Pandas. datanizing GmbH Mar 24 · 11 min read

The next set of two readings consider cleaning text and exploring the data. in the first post on cleaning text, simple regular expressions (regex) is used as well as advanced natural language processing tools. In the second post on exploratory data analysis for text, it explains loading data in order to explore data properties, summaries, frequencies, visualizations, and complexities.

3. How to solve 90% of NLP problems: a step-by-step guide: Using Machine Learning to understand and leverage text. Emmanuel Ameisen, Jan 24, 2018 · 13 min read

This overview reading presents the framework for solving natural language processing problems. It propose the following steps: 1) gather your data, 2) clean your data, 3) find a good data representation, 4) classification, 5) inspection, 6) Accounting for vocabulary structure, 7) Leveraging semantics, and 8) Leveraging syntax using end-to-end approaches.

[USEFUL OPTIONAL reading from Online Session 4 to get the big picture]

1. A Comprehensive Guide to Understand and Implement Text Classification in Python. SHIVAM BANSAL, APRIL 23, 2018

This comprehensive guide gives you all the necessary code to complete one iteration of the Machine Learning Life Cycle: 1) dataset preparation, 2) feature engineering, and 3) model building.

2. Machine Learning, NLP: Text Classification using scikit-learn, python and NLTK.
Javed Shaikh Jul 23, 2017 · 7 min read (beware of old deprecated code, but consider the example)

This reading demonstrates the python code in steps for one iteration of the machine learning life cycle for solving text classification. The python code includes 1) prerequisite and setting up the environment, 2) loading the data set into jupyter, 3) extracting features from text files, 4) running ML algorithms, 5) grid search for parameter tuning, and 6) useful tips and a touch of NLTK.

VIEW

1. Text Classification. Applied Text Mining in Python from University of Michigan, Course 4 of 5 in the Applied Data Science with Python Specialization [12:21 min]

The video introduces text classification, which also can be considered as supervised learning for text. It begins with a medical text example as well as several other examples. Then it formalizes the methodology and problem definition. Finally it presents the paradigms and questions to questions to ask when solving text classification.

2. Natural Language Processing, National Research University Higher School of Economics, Course 6 of 7 in the Advanced Machine Learning Specialization [14:30 min]

The next video goes through the text classification problem by considering the word (token). It also presents text cleaning by normalization and stemming with python.

CODING EXERCISE

From the readings of this lesson, there are complementary codes given in github, please only submit the two that are given below. Make sure you focus on the portion of the code for text pre-processing and data exploration; especially on topics such as exploring the data (including imbalanced classes), cleaning the text via simple regular expressions and advanced NLP, and exploring data properties, summaries, frequencies, visualizations, and complexities. It is alright if you do not understand all the steps in the machine learning life cycle right now, we will go over them in detail the up-coming weeks in this course. You will get practical hands-on experience by writing your own code.

After completing the programming code components, think about which methodologies you would employ in your project and how you would use them.

(The following links will open in new window)

1. Modern Text Mining with Python, Part 1 of 5: Introduction, cleaning and linguistics. datanizing GmbH Mar 24 · 7 min read
2. Modern Text Mining with Python, Part 2 of 5: Data Exploration with Pandas. datanizing GmbH Mar 24 · 11 min read.

Some Tips and Hints:

1. Some of the simple skills you will learn will be how to execute jupyter notebooks and install different libraries and modules. If you are using MAC and have have anaconda, a better way to install packages is to do it from the Anaconda environments UI. The order of installing packages is important so that they may not override other packages. You may have developed a checklist to try to get the code to begin working again. Next, please use virtual environments for each individual code exercise to organize the Python packages and their versions. You may look for tutorials online.

2. Another challenge you will face is dataset size and computational power. The dataset is too big on your machine to iterate over all the rows. You may reduce the dataset size to 50,000 rows by randomly selecting the dataset and balancing the classes. You may also run the code on the Google research collab environment so that you can see the results. The google environment provides an alternate environment for running commands against very large datasets. Although, this approach took a while to upload large datasets to the cloud environment. The last solution is to try GPU and parallel processing.

3. For "Multi-Class Text Classification with Scikit-Learn. Susan Li Feb 19, 2018 · 11 min read:-" The data file from :- https://catalog.data.gov/dataset/consumer-complaint-database is very big.

Int64Index: 463991 entries, 3 to 1437711

It takes a lot of time to compute. In order to decrease the processing time, change the starting lines to:-

```
import pandas as pd
df = pd.read_csv('Consumer_Complaints.csv')
df = df[:50000]
df.head()
```

Will change the data to :-

Int64Index: 15799 entries, 3 to 49999

DISCUSSIONS

Please post your weekly update in the appropriate thread in the discussion forum; DO NOT start your own thread. Remember to use the General Course Questions forum for any questions that you have about course content or logistics.

ASSIGNMENTS

Coding Exercise

Weekly coding exercise is submitted via PDF (please merge all the iPython exercises files together). Save your Jupyter notebooks for this week's Coding Exercise as one single PDF. Refer to Schedule for submission due dates.

Go to Assessments to Submit Online Session 1 - Coding Exercise

Discussions

Please post your weekly update in the appropriate thread in the discussion forum, please do not start your own thread. Remember to use the General Course Questions forum for any questions that you have about course content or logistics.

Project Proposal

Full details of what needs to be handed in can be found on the Course Outline Page (where the datasets are located). You can use the Project Discussion Forum to start working and collaborating with your classmates.

Supplementary Materials

Additional Reading Materials:

Text classification: the process of assigning tags or categories to text according to its content. It's one of the fundamental tasks in Natural Language Processing (NLP) with broad applications such as sentiment analysis, topic labeling, spam detection, and intent detection.

How to store the results in the SQLite database. How to drop the table -

Multiprocessing library

A paper discussing different sentiment analysis problems and challenges.

Additional Techniques for Handling Imbalanced Classes

For counting methods such as bag-of-words or tf-idf, methods like Tomek links, cluster centroids, and SMOTE will not work because they depend on realistic counts and co-dependence between the words (features.

Resampling strategies for imbalanced datasets [data here] - Rafael Alecar
This Jupyter notebook shows how to handle imbalanced data and use the evaluation metrics you just learned to assess the performance of imbalanced datasets and get to the best solution. First it introduces the imbalanced dataset and the metric trap from the confusion matrix. Then it introduces under and over sampling and the Python module. Lastly, it introduces some imbalanced learning methods such as tomek links, cluster centroids, and SMOTE.

Practical tips for class imbalance in binary classification?
This Python tutorial covers the example of class imbalance in binary classification. It discusses stratification (even sampling), performance metric, sampling for balance, and weighted (cost sensitive) learning

SUMMARY

Now that you have completed this module, you should know text classification and the type of real-world problems it solves. You should be able to begin solving a text classification problem, especially using them in your machine learning project.

You will continue to practice your coding skills in python and presenting your results in Jupyter notebook; these skills are crucial to the success of this course as well as your career. Don't forget to finish working on your project proposal.

Please work on the first online session this week and its coding exercise (due Sunday at 11:55 PM). Make sure to apply the techniques you learned on your project dataset. Lastly, remember to answer the 3 questions in your "Weekly Progress Update" in the Moodle Discussion Forum. Looking forward to seeing you again online.

Last modified: Friday, 17 April 2020, 1:52 PM

Midterm Review

MTE140
EN-SHIUN ANNIE LEE
9:30AM TUESDAY JUNE 08, 2010

Outline

- Pointer Example
- Linked List Example
- Recursion and Algorithm Complexity
- Big-Oh from first principles

Midterm Breakdown

Lecture Slides	Questions
Linked data representation	Q1
Recursion	Q2
Algorithm Complexity	Q2
List ADT	Q3
Stacks & Queues	Q4

What/How to Study?

1. Lecture notes
2. Programming – laboa, *labob, and project 1 (doublyLinkedList)
3. Previous midterm (but keeping in mind only what you have been taught this term, not last term)

LinkedList

```
// Functions Prototypes
void InsertNewLast( dataItem value, NodeType **L );
NodeType *ListSearch( dataItem value, NodeType *L );
```

LinkedList

```
int main()
{
  NodeType *head = NULL; // Declare list header as pointer
  InsertNewLast( 10, &head );
  InsertNewLast( 20, &head );
  InsertNewLast( 30, &head );

  // Searching...
  dataItem searchValue = 20;
  NodeType *nodePtr = ListSearch( searchValue, head );
  if (nodePtr)
       printf("Search value %i is FOUND in list\n", searchValue);
  else
       printf("Search value %i is NOT FOUND in list\n",
  searchValue);
}
```

LinkedList

```
// Functions Prototypes
void InsertNewLast( dataItem value, NodeType **L );
NodeType *ListSearch( dataItem value, NodeType *L );

int main()
{
NodeType *head = NULL; // Declare list header as pointer
InsertNewLast( 10, &head );
InsertNewLast( 20, &head );
InsertNewLast( 30, &head );

// Searching
dataItem searchValue = 20;
NodeType *nodePtr = ListSearch( searchValue, head );
if (nodePtr)
printf("Search value %i is FOUND in list/n", searchValue);
else
printf("Search value %i is NOT FOUND in list/n", searchValue);
}

typedef int itemType;
typedef struct listTag *List;

// Note that the structure definition and typedef can instead be combined into
// a single statement:
typedef struct node {
  itemType data;
  struct node *next;
  struct node *previous;
} nodeType;


struct listTag {
   nodeType *head;
   nodeType *tail;
   int size;
};
```

Recursion

```
Recursive Power(x, n)
//restrict n to be a power of 2
Int Power(int x, int n)
```

```
{
If(n==1 ) //base case
{

} else //recursive case\\
{
//return a smaller sub problem
}
}
```

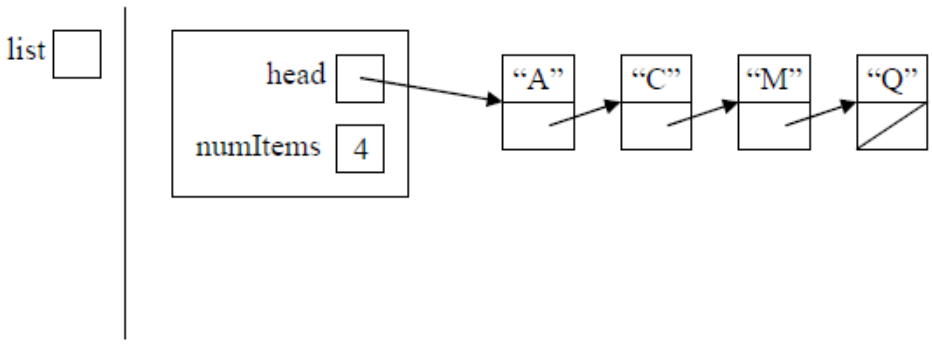
Big Oh from First Principles
A function $f(n)$ is said to be $O(g(n))$
if There exist 2 positive constants K and n_0
such that $|f(n)| \geq K * |g(n)| forall n \leq n_0$

Exam Question for MTE140 Algorithms and Data Structures

In the future I hope to teach the course described above, MTE140. The first year programming course for the one hundred or more mechatronics students is MTE140 "Algorithms and Data Structures". This course is taught once a year by a professor from the Systems Design Engineering department. Due to its infrequent offering by a different department, the course is not rigorously updated or refined. The coding exam questions included at the end of the response paper as "b." is taken from the exam. The code notation is taken from an optional lab 0b the students do not need to complete. Some effort is required to understand the underlying data structure and the function being used. Unfortunately, the solution code uses a specialized recursive call that students have not seen previously. This means students need specialized knowledge from outside the class and thus would not accurately evaluate the student's learning from the course. Because the question is not clearly defined and lacks details, a large variety of solutions can be given. This variety makes marking the question difficult, and the evaluation becomes focused on interpreting student's results rather than accessing their learning. The first portion of the course teaches two implementations of LinkedList and the ListADT. Below are three possible questions I would give to test these concepts.

Computational Question:

`a. (5 marks) Recall the doubly linked list implementation in Group Project 1, now assume you have a singly linked list that only has a pointer to the head of the list. The ListADT is as given below and a diagram of the singly linked list is given.`




```
typedef int itemType;
typedef struct listTag *List;

typedef struct node {
itemType data;
struct node *next;
} nodeType;

struct listTag {
nodeType *head;
int numItems;
};
```

Now write a function to print the list backwards (i.e., from last node element to the first node element) using recursion and returns a boolean value that checks whether the function is successfully called. Be sure to perform the appropriate error checking.

```
Bool PrintListBackwards(List L){
your base case code goes here
your recursive case code goes in here
}
```

b. (2 marks) Determine the recurrence relationship for the time required for PrintListBackwards in T(n).

c. (8 marks) Solve the recurrence relationship for the Merge sort algorithm using back substitution (unrolling), as well as give the function's order of growth in terms of the O notation.

The first computational question fully tests the student's knowledge of the material and ensures that the student is able to apply the coding skills they acquired from the project as well as the mathematic skills they acquired from attending the lecture. It assesses the student's ability to do the required work that was assigned in the hands-on programming projects. The student cannot guess the correct solution. Unfortunately, marking computational questions takes a long time. However with a clear description and some restriction to creative solutions it will hopefully reduce the marking effort.

Short Answer Question:

(2 marks) Give two reasons why you would use a circularly singly linked list instead of a doubly linked list. Note that for both lists you only have a pointer to the head.

The second short answer question is simple to answer and mark. There is some creativity in the answers the student can give as well as some subjectivity in the way the marker can mark. It also tests the basic understanding from the student simply.

Multiple Choice Question:
(1 mark) In group project one you implemented the DoublyLinkedLIst data structure and the ArrayList data structure. DoublyLinkedList is better than the ArrayList because
a) The traversal is faster when the data is sparse (correct solution)
b) the tail of the list is pointing to the last position (Distracter: true but unrelated to the stem, does not answer the question)
c) the doubly linked list needs to grow and double its size once it reaches its capacity (Distracter: false but related to the stem)
d) stacks and queues are equally efficient as lists (Distracter: false and unrelated to the stem)

The last multiple choice question is very easy to mark and its summary statistics can gage how well the class is learning. The stem of the question reminds the student of what they have learned in group project 1 and poses the question. It requires the student to think about each of the answers and apply what they have learned from group project 1. Note that each of the options and their functions are listed in brackets in the above question.

Appendix

Observations from Certificate in University Teaching program

Private & Confidential

Observation Report #2
En-Shiun Annie Lee – Systems Design Engineering

Event Observed:	SYDE 113 – Matrices and Linear Systems - TA
Date of Observation:	September 24th, 2010 at 2:30 – 3:20 PM
Location:	RCH 112
Students Present:	106 students
Prepared By:	Arash Shahi, TA Developer, Centre for Teaching Excellence

Context for Teaching Event: In this teaching event you were the TA for SYDE 133 (Matrices and Linear Systems), which is a mandatory first year course for students enrolled in Systems Design Engineering at the University of Waterloo. The tutorials for this course involved solving mathematical problems from the given optional problem sets. This particular tutorial was the first of the new semester and was therefore your first exposure to the new group of students. Your objectives for this review session were for your students to be able to convert between Cartesian and polar coordinates, plot Cartesian and polar coordinates, evaluate vector properties, and find complex conjugates and evaluate the functions.

Aspects to Maintain:

- **Delivery of Introduction and Rapport.** Annie, many elements of your tutorial introduction were excellent, as you were able to develop a rapport with the students early on. You appeared confident, personable, and enthusiastic, thus establishing yourself as an individual in addition to a qualified lecturer. Specifically, sitting upon the desk at the front of the class with your hands in your lap helped you seem relaxed and allowed you to demonstrate your care for student learning. You began the session with a brief self-introduction and explained why you were qualified to lead the tutorial. I particularly enjoyed when you explained your research in terms that the students could easily understand. You mentioned how your research involves DNA and bioinformatics, but utilizes large amounts of data "in a similar manner to Facebook, Twitter, etc.", which allowed you take something that first year students may not comprehend and relate it to something they can easily grasp. You then stated how you were interested in learning about the students within the tutorial and initiated a series of one-minute-paper activities to learn student names, something interesting about them, and what can be done to ensure they attend future tutorials. These activities, in addition to other elements of your introduction, clearly demonstrated your concern for student learning and helped establish a friendly rapport amongst yourself and the students.

- **Presentation Skills.** Annie, I am impressed to see that your presentation skills are consistently one of the strong aspects of your teaching. In both of your observations, you made your enthusiasm shine through your delivery. This way of presenting information is very helpful to students in a classroom because they learn to identify with the passion that you show toward the subject matter. I was particularly impressed with your enthusiasm early on in the tutorial. As you were not localized in front of the blackboard writing notes, you moved comfortably around the classroom ensuring that you were able to make eye contact

with a greater number of students. This was excellent because eye contact is an important aspect of interpersonal communication as it signals interest in others and encourages your students to participate. Your abundant use of eye contact early on in the tutorial also helped establish credibility, as it was your first tutorial session of the semester.

- **Structure & Classroom Mechanics.** Annie, after the initial introductions you began your lecture with an appropriate preview, outlining the topics you would cover. A preview is always a good idea, because it alerts your audience to your plan and mentally prepares them to digest what you have to say. It is especially helpful when your audience members need to take notes, because it helps them organize what they write. It was a good idea to keep the outline on the board for the entire class. Students responded very well to it and kept referring to it throughout the tutorial. During a longer presentation, you would want to return to this overview periodically to remind your audience what you have discussed and what you still plan to discuss in the rest of the talk.

- **Visual Aids.** Considering the fact that you were using only the blackboard to deliver visual aids, I thought that you were very able to succinctly convey the information that was necessary for your students' learning. Also, the structure of your board work was good as you progressively moved from the furthest left board to the furthest right board and used each separate board for independent topics. It made it very easy for students to take good notes and follow the tutorial. Finally, you were also able to use the blackboard efficiently when delivering important information about coordinate transformations. The graphs you drew on the board were particularly helpful for students.

- **One Minute Paper.** I was very impressed with your use of "one minute paper" in this tutorial. Even though I believe that it was a very aggressive move to incorporate four one minute papers in your first session with an undergraduate engineering class of over 100 students, it communicated many of your great personal and teaching related qualities to your students. On one paper, you asked them "tell me a little bit about yourself", which was a great way of creating a learning community and avoiding the "invisible student" way of thinking. You also asked them "what would motivate you to come to tutorials?", which in my opinion was a brilliant question. Not only you demonstrated that you care about their learning and would like them to be there, it also became clear that you are willing to go out of your way to motivate them and make the tutorials more useful for your students. I highly recommend that you limit the number of one minute papers to one per tutorial. With a class of that size a significant amount of time would be wasted by conducting an activity four times, while many benefits of this activity could have been captured with only one submission. Overall, It was a great experience and with a little more practice, I am sure you can fully utilize it in your classrooms. Knowing you through the last two observations, I understood very well how much you cared about your students and your enthusiasm towards teaching has come very clear, but I was very impressed with how you communicated that to the students on the very first tutorial! Great work!

Systems Design Engineering
Teaching Assistant Ranking Form

INSTRUCTOR: **Alex Wong**

Please rank your TAs and provide additional comments in the space below.
This information will be used for future selection of Teaching Assistantship assignments.

Student TA:	**En-Shiun Annie Lee**
Course Number:	**MTE 140**
Course Title:	**Mechatronics: Algorithms & Data Structure**
Term:	**Spring 2010**

Please Rank this TA by checking one of the following:

- ☑ **Excellent** - this is the highest recommendation for future Teaching Assistantships
- ☐ **Very Good** - would recommend again without hesitation
- ☐ **Sufficient** - acceptable and would work with TA again
- ☐ **Unacceptable** - would not work with TA again (please explain if areas for improvement were addressed, if improvements were noted etc.)

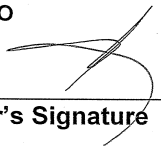
Additional Comments:

English Skills:
If the TA's first language is not English, please rate their English language skills:

☐ **Poor** ☐ **Average** ☐ **Good** ☑ **Excellent**

At the request of the student, would you authorize release of the information contained on this "TA Ranking Form"?

☑ **YES** ☐ **NO**



Instructor's Signature

Please return to Vicky Lawrence, DC 2641 asap. Thank you.

School of Computer Science

TEACHING ASSISTANT PERFORMANCE EVALUATION FORM

TA Supervisor: Leila Chinaei (ISC) **Term**: 1081 **Course:** CS 134 IA

Please rate the performance of the Teaching Assistant named below by entering a check-mark (√) in the appropriate box. Ignore any categories which are not applicable to your course. The **Teaching Assistant will receive a copy of this performance evaluation.** Please return this form to the mailbox in DC 2326A for Margaret Towell, Administrative Coordinator, Graduate Studies, as soon as possible. Thank you.

Teaching Assistant: Annie Lee

Rating Category	6	5	4	3	2	1
Industriousness	√					
Reliability & Responsibility	√					
Ability to Assist Students	√					
Preparation (before labs, etc.)	√					
Knowledge of Material	√					
Punctuality (lab attendance, returning marked papers, etc.)		√				
English Language Fluency	√					
Overall Rating	√					

Rating Scale:
6 = Excellent; 5 = Very Good; 4 = Acceptable;
3 = Below Average; 2 = Unacceptable; 1 = Insufficient Opportunity to Observe.

How many hours per week (on average) were required? 5

How many weeks did the student perform the TA duties? 16

Please provide a brief overview of duties performed by the TA:

leading tutorials, proctoring & marking midterm & final exam

Remarks: Very interested in the job

If you would like this student considered for a TA award, please fill out a nomination form.

Signature: Eduehinaei **Date:** April 30, 08

UNIVERSITY OF WATERLOO

ASSOCIATE PROVOST, GRADUATE STUDIES
200 University Avenue West, Waterloo, ON, Canada N2L 3G1
519-888-4567, ext. 33439 | fax 519-746-3051 | grad.uwaterloo.ca

March 29, 2011

En-Shiun (Annie) Lee
Systems Design Engineering
DC

Dear Ms. Lee:

By now you may be aware that the recipients of the Amit and Meena Chakma Award for Exceptional Teaching by a Student for 2011 were announced in Senate at the March 28, 2011 meeting. I would like to inform you that you were one of those considered for the Award.

The Selection Committee considered your nomination with great care and was impressed with the many positive comments made about your teaching by your students and professors. Clearly, your teaching is greatly appreciated. The Committee has asked me to convey the high regard for your teaching reflected in your nomination.

I would also like to extend to you my own sincere thanks for your commitment to teaching at the University of Waterloo. Please join us at the "Teaching Excellence Celebration" event being held Wednesday, May 11, 2011 at the University Club from 3:30 pm to 5:00 pm.

Yours sincerely,

Susan Horton
Chair, AETS Selection Committee

cc: F. Hamdullahpur, President
D. Ellis, Director, CTE
A. Sedra, Dean, Faculty of Engineering
P. Fieguth, Chair, Department of Systems Design Engineering

CAUCE
Canadian Association for University Continuing Education

CAUCE Secretariat
477 - 221 Cumberland Avenue North
Saskatoon, Saskatchewan
Canada S7N 1M3

Tel: 1-306-966-5604
Fax: 1-306-966-5590
cauce.secretariat@usask.ca
cauce-aepuc.ca

March 31, 2020

Sean Woodhead, Program Manager
School of Continuing Studies
York University
96 The Pond Road
Toronto Ontario M3J 1P3

Dear Sean:

I am pleased to announce that your submission to the 2020 CAUCE Program Awards competition, *Certificate in Machine Learning* has been selected for the award in the *Non-credit Programming over 48 hours* category. On behalf of the CAUCE Executive and the Program Awards Committee, I would like to extend congratulations to you and your colleagues.

To assist with the preparation of your award certificate and ensure accuracy, please provide the secretariat with the information requested on the enclosed sheet **before April 30th.** Also, if you would like the incoming president of CAUCE, Ian Allen, to send a letter to your university president (with a copy to your dean or director) advising them of the award, please include that information on the form as well.

Because of the cancellation of our annual conference in 2020 due to COVID-19, your award certificate will be mailed to you this spring. The secretariat will confirm the correct mailing address with you at that time.

Once again, my congratulations on winning a 2020 CAUCE 2020 award and I am sorry that we will not be able to celebrate in person this year.

Sincerely,

Heather McRae, CAUCE President

cc Jo-Anne Clarke, Chair, Program Awards Committee
Christie Schultz, CAUCE Executive Liaison to the Program Awards Committees