AML-2304 Natural Language Processing

Computer Studies

Course Number:	Co-Requisites:	Pre-Requisites:
AML-2304	N/A	N/A
Prepared by:	Lambton College, Outline Creator	
Approved by:	Chris Slade, Dean International Education	
Approval Date:	Friday, June 4, 2021	
Approved for Academic Year:	2021-2022	
Credit Weight:	4.00	

Course Description

This course introduces Natural Language Processing (NLP) and its key concepts. The theory part includes the use of classic machine learning methods to solve machine translation, language modeling, and sequence tagging. The laboratory portion is designed to provide students with the opportunity to work with a set of NLP problems and the opportunity to apply their knowledge to resolve them.

Course Learning Outcomes/Course Objectives

1. **Evaluate the key concept of Natural Language Processing (NLP)**
   - 1.1 Define Natural Language Processing
   - 1.2 Describe the Challenges of NLP; extraction of meaning from text
   - 1.3 Discuss the importance of Natural Language Processing

2. **Evaluate the application of NLP in various Canadian businesses**
   - 2.1 Discuss the usage of NLP in Customer Service
   - 2.2 Describe how NML is used in Market Intelligence
   - 2.3 Discuss the applicability of NLP for Sentiment Analysis

3. **Explain Text Wrangling and Cleansing**
   - 3.1 Discuss Sentence Splitting
   - 3.2 Define Word Tokenization
   - 3.3 Explain Stemming and Lemmatization
   - 3.4 Discuss Stop Words

3.5 Discuss Regular Expression (Regex)

3.6 Explain Bag-of-Words

3.7 Discuss TF-IDF

4. Demonstrate Text Preprocessing: Replacing and Correcting Words

4.1 Illustrate Text conversion to Lowercase

4.2 Apply Number Removal, Punctuation Removal, and Whitespace Removal

4.3 Use Parts of Speech Tagging (POS)

4.4 Practice Named Entity Recognition

4.5 Show Collocation Extraction and Synonyms

4.6 Demonstrate use of Word Lengthening to detect sentiment

4.7 Apply Spell Correction

5. Illustrate Text Classification

5.1 Import Libraries and the dataset

5.2 Perform Text Preprocessing

5.3 Perform feature extraction from Textual data

5.4 Prepare Training and Test Sets by splitting dataset

5.5 Demonstrate The training of the Text Classification Model (Naïve Bayes)

5.6 Analyze the results on test set and Evaluate the Model

5.7 Demonstrate saving and loading the Model

6. Evaluate Sentiment Analysis

6.1 Discuss VADER Sentiment Analysis

6.2 Discuss working and scoring

6.3 Discuss Punctuation, Capitalization, Degree Modifier, and Conjunctions

6.4 Explain Preceding Trigram

7. Discriminate Parsing Structure in Text

7.1 Describe Chunking (Shallow Parsing)

7.2 Discuss Chinking as an extension of Chunking

8. Evaluate Sentiment Analysis using Python NTLK (Natural Language Toolkit) library

8.1 Describe Supervised Classification

8.2 Construct a List of Movie Review Document

8.3 Perform Feature extraction

8.4 Remove Punctuation and Stop-words

8.5 Compute Frequency Distribution of Cleaned Word List

8.6 Create a Word Feature using most frequently occurring words

8.7 Construct a Feature Set

8.8 Apply training for the Classifier

8.9 Show and explain the result of testing

9. Compare Text Classification using Scikit-Learn

9.1 Import textual datasets into Python

9.2 Extract features from text files

9.3 Compare ML Classifier Algorithms [Naïve Bayes, Support Vector Machine (SVM) etc.]

10. Demonstrate Files operations in Python

10.1 Illustrate reading and writing to Text files

10.2 Show reading and to PDF files

11. Evaluate Word2Vec Algorithm

11.1 Discuss Word2Vec Algorithm

11.2 Describe Neural Word Embeddings

11.3 Explore Advances in NLP: ElMO, BERT and GPT-2

11.4 Show Word2vec Use Cases

Relationship to Vocational Learning Outcomes

This course provides the opportunity for you to achieve the following Program Vocational Learning Outcomes (VLO) which will be taught and evaluated at an taught (T), assessed (A) or culminating performance (CP) level:

AIMT - Artificial Intelligence & Machine Learning

VLO 1	Collect, manipulate and mine data sets to meet organizational need. (T, A)
VLO 3	Design and apply data models that meet the needs of a specific operational/business process. (T, A)
VLO 7	Identify and assess data analytics business strategies and workflows to respond to new opportunities or provide project solutions. (T, A)
VLO 9	Select and apply appropriate artificial intelligence and machine learning techniques and algorithms that meet the needs of a specific operational/business process. (T, A)

DSMM - Big Data Analytics

VLO 2	Recommend different systems and network architectures, artificial intelligence and data storage technologies to support data analytics and Big Data. (T, A)
VLO 4	Develop software applications, algorithms and artificial intelligence models to manipulate, correlate and reduce data sets and produce project documentation and reports. (T, A)

VLO 10	Develop artificial intelligence solutions to support administration, decision-making, planning, risk management, logistics, manufacturing, smart devices and robotics. (T, A)

Learning Resources

Required:

Natural Language Processing: A Quick Introduction to NLP with Python and NLTK
Samuel Burns
ISBN-13: 978-1079244311
155 pages
Release Date: 8 July 2019
Published by: Independently published

Supplemental:

- Personal Computer

Student Evaluation

Tests (40%):

- Midterm Test (20%)
- Final Exam (20%)

Assignments and Projects (60%):

- Assignment (20%)
- Final Term Project (40%)

Grade Scheme

The round off mathematical principle will be used. Percentages are converted to letter grades and grade points as follows:

Mark (%)	Grade	Grade Point	Mark (%)	Grade	Grade Point
94-100	A+	4.0	67-69	C+	2.3
87-93	A	3.7	63-66	C	2.0
80-86	A-	3.5	60-62	C-	1.7
77-79	B+	3.2	50-59	D	1.0
73-76	B	3.0	0-49	F	0.0
70-72	B-	2.7			

Prior Learning Assessment and Recognition

Students who wish to apply for prior learning assessment and recognition (PLAR) need to demonstrate competency at a post-secondary level in all of the course learning requirements outlined above. Evidence of learning achievement for PLAR candidates includes:

- Not Applicable: Post-graduate course

Course Related Information

This is a project and discussion-based learning course consisting of a variety of teaching environments. All course work is completed according to the syllabus. Students should take careful notes as not all material can be found in the textbook or handout materials. Attendance is expected and necessary to be successful.

College Related Information

Academic Integrity

Lambton College is committed to high ethical standards in all academic activities within the College, including research, reporting and learning assessment (e.g. tests, lab reports, essays).
The cornerstone of academic integrity and professional reputation is principled conduct. All scholastic and academic activity must be free of all forms of academic dishonesty, including copying, plagiarism and cheating.

Lambton College will not tolerate any academic dishonesty, a position reflected in Lambton College policies. Students should be familiar with the Students Rights and Responsibilities Policy, located at lambtoncollege.ca. The policy states details concerning academic dishonesty and the penalties for dishonesty and unethical conduct.

Questions regarding this policy, or requests for additional clarification, should be directed to the Lambton College Student Success Department.

Students with Disabilities

If you are a student with a disability please identify your needs to the professor and/or the Accessibility Centre so that support services can be arranged for you. You can do this by making an appointment at the Accessibility Centre or by arranging a personal interview with the professor to discuss your needs.

Student Rights and Responsibility Policy

Acceptable behaviour in class is established by the instructor and is expected of all students. Any form of misbehaviour, harassment or violence will not be tolerated. Action will be taken as outlined in Lambton College policy.

Date of Withdrawal without Academic Penalty

Please consult the Academic Regulations and Registrar's published dates.

Waiver of Responsibility

Every attempt has been made to ensure the accuracy of this information as of the date of publication. The content may be modified, without notice, as deemed appropriate by the College.

Students should note policies may differ depending on the location of course offering. Please refer to campus location specific policies:
LAMBTON COLLEGE POLICIES – applicable to all Lambton College students.
- Student Rights & Responsibilities & Discipline policy (2000-5-1)
- Test & Exam Writing Protocol (2000-1-6)
- Evaluation of Students (2000-1-3)
- (https://www.lambtoncollege.ca/custom/Pages/Policies/Policies.aspx)

CESTAR COLLEGE:
- https://www.lambtoncollege.ca/Programs/International/Lambton_in_Toronto/Student_Policies/

QUEENS COLLEGE:
- https://www.lambtoncollege.ca/Programs/International/Lambton_in_Mississauga/Student_Policies/

Note: It is the student’s responsibility to retain course outlines for possible future use to support applications for transfer of credit to other educational institutions.