Cover Sheet: Request 14841

Natural Language Processing Course in CISE

Info

Process	Course|New|Ugrad/Pro
Status	Pending at PV - University Curriculum Committee (UCC)
Submitter	Peter Dobbins ichabod@ufl.edu
Created	3/30/2020 8:34:12 PM
Updated	9/15/2020 4:05:19 PM
Description of request	This is request for a new course prefix and number to be created so that the Computer & Information Science & Engineering (CISE) department can introduce a formal course on Natural Language Processing (NLP) to the undergraduate students at UF.

Actions

Step	Status	Group	User	Comment	Updated
Department	Approved	ENG - Computer and Information Science and Engineering 011914001	Arunava Banerjee		6/2/2020
No document changes					
College	Approved	ENG - College of Engineering	Heidi Dublin	Approved by HWCOE Curriculum Committee and Faculty Council.	9/15/2020
syllabus_nlp.pdf					9/9/2020
University Curriculum Committee	Pending	PV - University Curriculum Committee (UCC)			9/15/2020
No document changes					
Statewide Course Numbering System					
No document changes					
Office of the Registrar					
No document changes					
Student Academic Support System					
No document changes					
Catalog					
No document changes					
College Notified					
No document changes					

Course|New for request 14841

Info

Request: Natural Language Processing Course in CISE
Description of request: This is request for a new course prefix and number to be created so that the Computer & Information Science & Engineering (CISE) department can introduce a formal course on Natural Language Processing (NLP) to the undergraduate students at UF.
Submitter: Peter Dobbins ichabod@ufl.edu
Created: 10/12/2020 1:37:58 PM
Form version: 3

Responses

Recommended Prefix CAP
Course Level 4
Course Number XXX
Category of Instruction Intermediate
Lab Code None
Course Title Natural Language Processing
Transcript Title Natural Language Processing
Degree Type Baccalaureate

Delivery Method(s) On-Campus
Co-Listing No

Effective Term Earliest Available
Effective Year Earliest Available
Rotating Topic? No
Repeatable Credit? No

Amount of Credit 3

S/U Only? No
Contact Type Regularly Scheduled
Weekly Contact Hours 3
Course Description Introduction to the essential concepts, principles, and techniques of Natural Language Processing (NLP). Practical application and theoretical concepts are examined. Topics include information extraction, language construction, grammars, disambiguation, as well as system modeling, classification, and evaluation.
Prerequisites CAP 3530
Co-requisites N/A
Rationale and Placement in Curriculum Natural Language Processing (NLP) is a significant area of Machine Learning, having become its own field of study. We see possibility of many practical applications for NLP in social media, ranging from speech recognition assistance to flagging offensive material. Students with NLP experience will be able to utilize NLP expertise within other related fields of study, including Bioinformatics, Data Science, Human Centered Computing (HCC), and Human Computer Interaction (HCI). All of which are already taught here at UF. However, UF does not have an undergraduate level NLP course.
Course Objectives Students will be able to apply their knowledge to:
* evaluate the performance of natural language applications
* search abstract data sets (for example the deep-web) for query answer retrieval
* develop, implement, and evaluate their own programmatic solutions to natural language problems
* apply standard mathematical and statistical principles to natural language problems
Course Textbook(s) and/or Other Assigned Reading The Second edition of Natural Language Processing with Python, currently published for free here: http://www.nltk.org/book/.

For reference, here is the first edition
* Natural Language Processing with Python, First Edition, Bird, Klein, & Loper, 2009, O'Reilly Media, ISBN 0-596-51649-5.

Weekly Schedule of Topics The following data is based upon 43 lecture periods over approximately 15, pending Holidays and UF scheduling dates. A pdf / docx copy is also available if easier to read.

Week #1
Reading: Syllabus, Ch 1.1 - 1.2
Lecture
1: Syllabus, Course Policies
2: Introduction to Information Extraction
3: Installation and Using Python
HW #1 Assigned

Week #2
Reading: Ch 1.3 - 1.8; 2.1
Lecture
4: Bigrams, Collocations, NLTK Functions
5: Python Functions, Frequency Distribution, Loops
6: Testing vs Training, Lists, Overfitting, Corpora
HW #1 Due
HW #2 Assigned

Week #3
Reading: Ch 2.1 - 2.8
Lecture
7: Conditional Frequency Distribution
8: Classifying Antagonizing Text, Wordnet
9: Similarity Measure, Grammatical Definitions
HW #2 Due
HW #3 Assigned

Week #4
Reading: N/A
Lecture
10: Zipf's Law, Exam I Review
11: Exam I
12: File IO
HW #3 Due

Week #5
Reading: Ch 3.1 - 3.12
Lecture
13: File IO
14: Regular Expressions
15: Regular Expressions
HW #4 Assigned

Week #6
Reading: Ch 4.1 - 4.11
Lecture
16: Problem Solving in Python
17: Problem Solving in Python
18: Sample Exercises
HW #4 Due
HW #5 Assigned

Week #7
Reading: N/A
Lecture
19: Exam II Review
20: Exam II
21: Finite State Automata
HW #5 Due

Week #8
Reading: Ch 5.1 - 5.5
Lecture
22: Collocations, Parts of Speech
23: Variance, t-Test
24: Mean Differential, Chi-Square Test
HW #6 Assigned

Week #9
Reading: Ch 5.5 - 5.10
Lecture
25: Mean Differential, Chi-Square Test
26: Parts of Speech
27: Parts of Speech
HW #6 Due
HW #7 Assigned

Week #10
Reading: N/A
Lecture
28: Exam III Review
29: Exam III
30: Tagging
HW #7 Due

Week #11
Reading: Ch 6.1 - 6.10
Lecture
31: Confusion Matrix, Precision, Recall
32: Cross Validation
33: Supervised Classification
HW #8 Assigned

Week #12
Reading: Ch 7.1 - 7.9
Lecture
34: Bayesian Analysis
35: Feature Extraction
36: Classifiers, Decision Trees, Max Entropy
HW #8 Due
HW #9 Assigned

Week #13
Reading: N/A
Lecture
37: Exam IV Review
38: Exam IV

39: Chunking
HW #9 Due

Week #14
Reading: Ch 8.1 - 8.9
Lecture
40: Context Free Grammars
41: Context Free Grammars
42: Exam V Review
HW #10 Assigned

Week #15
Reading: N/A
Lecture
43: Exam V
44: N/A
45: N/A
HW #10 Due

Week #16: Final Exams Week
Reading: N/A
Lecture: N/A
HW #1 N/A
Grading Scheme 5 exams worth 10% each for a total of 50% of the grade. Exams will test students on material since the previous exam. Problems will include definitions, mathematical, and programmatic. Instructor solutions will identify key elements required and grades will be assessed based upon the presence of these elements in the student solution.

10 homeworks [1 drop of lowest score] worth 50% of the grade [5.55% each]. Exercises will include programming projects and statistical evaluation. Grades will be based upon the execution of programs. In the event of incomplete student work, instructor solutions will identify key elements required and grades will be assessed based upon the presence of these elements in the student solution.

Grading Scale
A 92 to 100
A- 89.00 to 91.99
B+ 86.00 to 88.99
B 82.00 to 85.99
B- 79.00 to 81.99
C+ 76.00 to 78.99
C 72.00 to 75.99
C- 69.00 to 71.99
D+ 66.00 to 68.99
D 62.00 to 65.99
D- 59.00 to 61.99
E 0.00 to 58.99
Instructor(s) Peter J Dobbins
Attendance & Make-up Yes
Accomodations Yes
UF Grading Policies for assigning Grade Points Yes
Course Evaluation Policy Yes

Couldn't create PDF for syllabus_nlp.pdf

Download PDF here

CAP 4XXX Natural Language Processing

Course Prefix / Number Sections: TBD
Class Periods: TBD
Location: TBD
Academic Term: TBD

Instructor:
Pete Dobbins
pjd@cise.ufl.edu
352.294.6685
Office Hours: M 8th period [15:00 – 15:50], WR 4th period [10:40 – 11:30] in CSE E474

Course Description
Introduction to the essential concepts, principles, and techniques of Natural Language Processing (NLP). Practical application and theoretical concepts are examined. Topics include information extraction, language construction, grammars, disambiguation, as well as system modeling, classification, and evaluation.

Course Pre-Requisites / Co-Requisites
Prerequisite: COP 3530

Course Objectives
Students will learn the practical application and theory of Natural Language Processing. Practically, we are motivated to study natural language to better understand how to search abstract data sets (for example the *deep-web*) for query answer retrieval. Among the theoretical topics included are: regular expressions, collocations, n-grams, syntactic parsing, word sense disambiguation, lexical acquisition, Markov models, context-free grammars, and information retrieval. **Python** will be introduced and used to illustrate course concepts. Project solutions will be developed in **Python**. Weekly course topics are provided on the calendar pages of the course website.

Students will be able to apply their knowledge to:
- evaluate the performance of natural language applications
- search abstract data sets (for example the deep-web) for query answer retrieval
- develop, implement, and evaluate their own programmatic solutions to natural language problems
- apply standard mathematical and statistical principles to natural language problems

Materials and Supply Fees
There are no materials or supply fees for this course beyond the textbook.

Required Textbooks and Software
- The second edition of *Natural Language Processing with Python* is published here: http://www.nltk.org/book/.

Supplemental Textbooks
- *Hands-on Machine Learning with Scikit-Learn & TensorFlow*, Second Edition, Aurelien Geron, O'Reilly Media, 2019, ISBN 1-492-03264-6
- *Learning Python*, Fourth Edition, Mark Lutz, 2009, O'Reilly Media, ISBN 0-596-15806-8.
- *Foundations of Statistical Natural Language Processing*, First Edition, Manning & Schuetze, 1999, ISBN 0-262-13360-1.
- *Speech and Language Processing*, Second Edition, Jurafsky & Martin, 2009, Prentice Hall, ISBN 0-131-87321-0.

Resources
Recommendations for other resources are given here.

- *Natural Language Processing with Python*, First Edition, Bird, Klein, & Loper, 2009, O'Reilly Media, ISBN 0-596-51649-5. (this is the first edition of the textbook; it is not updated for NLTK 3 and Python 3)
- *Errata* from the June 2009 printing: http://nltk.googlecode.com/svn/trunk/nltk/doc/book/errata.txt.
- Website: http://www.nltk.org/.
- Python: http://www.python.org/.
- Learn and Practice Python: http://inventwithpython.com/.
- NLP on Wikipedia: http://en.wikipedia.org/wiki/Natural_language_processing.
- Association for the Advancement of Artificial Intelligence (AAAI): http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/NaturalLanguage.

Course Schedule

Here is a week-by-week schedule for a 43 lecture semester. Lecture numbers will remain the same while dates will adapt to the given semester and University holiday schedule. Please refer to the announcements and Canvas Modules for additional details and material regarding each topic.

Week #	Homework	Lecture #	Topics	Reading
1	1 assigned	1	Syllabus, Course Policies	Syllabus 1.1 – 1.2
		2	Introduction to Information Integration	
		3	Installation and Using Python	
2	1 due 2 assigned	4	Bigrams, Collocations, NLTK Functions	1.3 – 1.8 2.1
		5	Python Functions, Frequency Distribution, Loops	
		6	Testing vs Training, Lists, Overfitting, Corpora	
3	2 due 3 assigned	7	Conditional Frequency Distribution	2.1 – 2.8
		8	Classifying Antagonizing Text, Wordnet	
		9	Similarity Measures, Grammatical Definitions	
4	3 due	10	Zipf's Law, Exam I Review	
		11	Exam I	
		12	File IO	
5	4 assigned	13	File IO	3.1 – 3.12
		14	Regular Expressions	
		15	Regular Expressions	
6	4 due 5 assigned	16	Problem Solving in Python	4.1 – 4.11
		17	Problem Solving in Python	
		18	Sample Exercises	
7	5 due	19	Exam II Review	
		20	Exam II	
		21	Finite State Automata	
8	6 assigned	22	Collocations, Parts of Speech	5.1 – 5.5
		23	Variance, t Test	
		24	Mean Differential, Chi-Square Test	
9	6 due 7 assigned	25	Mean Differential, Chi-Square Test	5.5 – 5.10
		26	Parts of Speech	
		27	Parts of Speech	
10	7 due	28	Exam III Review	
		29	Exam III	
		30	Tagging	
11	8 assigned	31	Confusion Matrix, Precision, Recall	6.1 – 6.10
		32	Cross Validation	
		33	Supervised Classification	
12	8 due 9 assigned	34	Bayesian Analysis	7.1 – 7.9
		35	Feature Extraction	
		36	Classifiers: Decision Tree, Max Entropy	

13	9 due	37	Exam IV Review	
		38	Exam IV	
		39	Chunking	
14	10 assigned	40	Context Free Grammars	8.1 – 8.9
		41	Context Free Grammars	
		42	Exam V Review	
15	10 due	43	Exam V	
16	N/A	N/A	Final Exams Week	N/A

Attendance Policy, Class Expectations, and Make-Up Policy

General policies can be found here: http://www.cise.ufl.edu/~pjd/admin/policies/course.html. You are expected **know and follow** all of these policies.

In order to be successful in this course, you should attend lecture. There is no replacement to the experience of seeing the material presented during lecture.

Make ups for graded class activities are provided given appropriate documentation is presented. Excused absences must be consistent with University policies in the undergraduate catalog (https://catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/) and meet University requirements regarding excused absences. There are very few events which would impede your participation in a graded activity warranting an exception, that you would not be aware of well before [at least a few hours] the start of the activity.

Late work [such as *homework*, this is not applicable to in class activities or Exams which must be completed by the assigned date/time of the activity unless the absence is excused]: No work can be accepted after a solution is posted. If a solution has not been posted and you submit one day late, there will be a penalty of up to 50% on the graded work. No work that is two days late can be accepted. Note, if the assignment is due at 11:59 PM, then at 12:00 AM the work is one day late.

Evaluation of Grades

Activity	Final Grade Percentage
Exams [5]	50% [10% each]
Homework	50%
Total	100%

Exams

The exam schedule is provided here. There will be five exams. All exams will be given in lecture. Each exam is worth 10% of your final grade.

Exam	Day	Week #
I	Wednesday	4
II	Wednesday	7
III	Wednesday	10
IV	Wednesday	13
V	Monday	15

Homework (AKA Project Assignments)

Project assignments will include written homework problems and the implementation of Python programs. Your *lowest* project score will be *dropped*. The details of project submissions will be given for each assignment and include submissions given by hard copy in lecture and soft copy via e-Learning (Canvas) by 11:59 PM on the due date assigned within the project specification. If you will be absent from lecture, ensure any assigned hard copies have been delivered to before the lecture meeting time.

Practice Exercises

There will be practice exercises assigned during lecture that are not associated with a graded homework. These exercises are supplemental to the course lectures and graded materials, helping you in your studies and preparation. These exercises will not be collected, however if you would like to discuss your solutions with us, please do so. It is *highly* recommended that you complete and participate in non-graded course activities: *lecture* and *practice exercises*, in addition to the graded activities: *exams* and *homework*.

Grading Policy

The range used to calculate your final letter grade in our course will be no harsher than this grading scale provided in the following table. Your final point total will be calculated using the percentages given in the *Evaluation of Grades* section. The percent you earn on each activity will be multiplied by the grade points associated with that activity.

Grade Points	Letter Grade	Highest	Lowest
4.00	A	100+	92.00
3.67	A–	91.99	89.00
3.33	B+	88.99	86.00
3.00	B	85.99	82.00
2.67	B–	81.99	79.00
2.33	C+	78.99	76.00
2.00	C	75.99	72.00
1.67	C–	71.99	69.00
1.33	D+	68.99	66.00
1.00	D	65.99	62.00
0.67	D–	61.99	59.00
0.00	E	58.99	0.00

More information on UF grading policy may be found at:
https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/

Re-Grade Requests

You must return exams/projects/quizzes to the person grading the activity in question, along with a **printed (not handwritten) attachment** stating the reasons for the re-grade request. Do not write anything on the exam/project/quiz itself or detach (un-staple) the pages from each other. If you do so, your exam/project/quiz will not be re-graded. Note, when requesting a re-grade, the whole exam/project/quiz is subject to the re-grade, not just the portion you specify.

Special Notes for the First Week of Classes

There will be no office hours the first week of classes. If you need to see me, do so after lecture or schedule an individual appointment.

Students Requiring Accommodations

Students with disabilities who experience learning barriers and would like to request accommodations should connect with the Disability Resource Center https://disability.ufl.edu/students/get-started/ (352.392.8565, https://disability.ufl.edu/). Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester, delaying may affect the ability to schedule the accommodation. All exams, quizzes, and

so on administered through the DRC office must follow the course exam structure, being completed on the same date and time as the course exam.

Course Evaluation
Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

University Honesty Policy
UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (https://sccr.dso.ufl.edu/policies/student-honor-code-student-conduct-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

Commitment to a Safe and Inclusive Learning Environment
The Herbert Wertheim College of Engineering values broad diversity within our community and is committed to individual and group empowerment, inclusion, and the elimination of discrimination. It is expected that every person in this class will treat one another with dignity and respect regardless of gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture.

If you feel like your performance in class is being impacted by discrimination or harassment of any kind, please contact your instructor or any of the following:
- Your academic advisor or Graduate Program Coordinator
- Robin Bielling, Director of Human Resources, 352-392-0903, rbielling@eng.ufl.edu
- Curtis Taylor, Associate Dean of Student Affairs, 352-392-2177, taylor@eng.ufl.edu
- Toshikazu Nishida, Associate Dean of Academic Affairs, 352-392-0943, nishida@eng.ufl.edu

Software Use
All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

Student Privacy
There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see: https://registrar.ufl.edu/ferpa.html

Campus Resources:

Health and Wellness

U Matter, We Care:
Your well-being is important to the University of Florida. The U Matter, We Care initiative is committed to creating a culture of care on our campus by encouraging members of our community to look out for one another and to reach out for help if a member of our community is in need. If you or a friend is in distress, please contact umatter@ufl.edu so that the U Matter, We Care Team can reach out to the student in distress. A nighttime and

weekend crisis counselor is available by phone at 352-392-1575. The U Matter, We Care Team can help connect students to the many other helping resources available including, but not limited to, Victim Advocates, Housing staff, and the Counseling and Wellness Center. Please remember that asking for help is a sign of strength. In case of emergency, call 9-1-1.

Counseling and Wellness Center: http://www.counseling.ufl.edu/cwc, and 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

Sexual Discrimination, Harassment, Assault, or Violence
If you or a friend has been subjected to sexual discrimination, sexual harassment, sexual assault, or violence contact the **Office of Title IX Compliance**, located at Yon Hall Room 427, 1908 Stadium Road, (352) 273-1094, title-ix@ufl.edu

Sexual Assault Recovery Services (SARS)
Student Health Care Center, 392-1161.

University Police Department at 392-1111 (or 9-1-1 for emergencies), or http://www.police.ufl.edu/.

Academic Resources

E-learning technical suppor*t*, 352-392-4357 (select option 2) or e-mail to Learning-support@ufl.edu. https://lss.at.ufl.edu/help.shtml.

Career Resource Center, Reitz Union, 392-1601. Career assistance and counseling. https://www.crc.ufl.edu/.

Library Support, http://cms.uflib.ufl.edu/ask. Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center, Broward Hall, 392-2010 or 392-6420. General study skills and tutoring. https://teachingcenter.ufl.edu/.

Writing Studio, 302 Tigert Hall*,* 846-1138. Help brainstorming, formatting, and writing papers. https://writing.ufl.edu/writing-studio/.

Student Complaints Campus*:* https://www.dso.ufl.edu/documents/UF_Complaints_policy.pdf.

On-Line Students Complaints*:* http://www.distance.ufl.edu/student-complaint-process.