1 Course Information

1.1 Meetings and Instructors

Monday, Wednesday, 1:00 – 2:15 PM
Location: Sondheim, 101

https://www.csee.umbc.edu/courses/undergraduate/473/f19/

This class will use the Piazza discussion forum

http://piazza.com/umbc/fall2019/cmsc473673/home

where course announcements and extracurricular questions and discussions can happen. Each person is responsible for staying up-to-date on the Piazza posts.

For any sensitive issue, please email me (ferraro@umbc.edu).

Prerequisites CMSC 341 (Principles of Programming Languages) and one of {STAT 355, CMPE 220, STAT 451}, all with the grade of C or better; or instructor permission
## 1.2 Texts, Readings, and Discussion

The readings will be from publicly available sources. Most will come from the (draft) of the third edition of Dan Jurafsky and James H. Martin’s *Speech and Language Processing*, available at [https://web.stanford.edu/~jurafsky/slp3/](https://web.stanford.edu/~jurafsky/slp3/) Other readings will come from papers published in the ACL community, available at [http://aclanthology.info](http://aclanthology.info).

The second edition of Jurafsky and Martin is a recommended, but not required, text. This is available in the library. Readings from the 2nd edition can often be substituted for the 3rd. (Another classic text is the 1999 *Foundations of Statistical Natural Language Processing*, by Manning and Schütze.)

## 1.3 Topics

The topics covered will range from basic counting, probability and classification, through language modeling (n-gram models, smoothing heuristics, maxent/log-linear models, and distributed/vector-valued representations), to sequences of latent variables (e.g., hidden Markov models, some basic machine translation alignment), to trees, graphs (syntax and semantics), some discourse-related applications (coreference resolution, RTE, etc.), and other special or current topics (e.g., fairness and ethics in NLP).

## 1.4 Goals

After taking this course, you will:

1. be introduced to some of the core problems and solutions of NLP;
2. learn different ways that success and progress can be measured in NLP;
3. be exposed to how these problems relate to those in statistics, machine learning, and linguistics;
4. have experience implementing a number of NLP programs;
5. read and analyze research papers;
6. practice your (written) communication skills.
2 Coursework and Evaluation

The undergraduate version consists of five or six assignments, a midterm, and a course project. The graduate version is everything in the undergraduate version, plus a synthesis paper that must undergo substantial revision. The overall evaluation is as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>473—Undergrad</th>
<th>673—Grad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignments</td>
<td>45%</td>
<td>30%</td>
</tr>
<tr>
<td>Midterm</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Graduate Paper</td>
<td>–</td>
<td>30%</td>
</tr>
<tr>
<td>Project</td>
<td>45%</td>
<td>30%</td>
</tr>
</tbody>
</table>

**Intracomponent Weighting** A component’s score will be the maximum between the macro-average and the micro-average of the individual assessments within that component. If your assignment scores are 65/90, 95/100, 95/110, and 100/110, then the micro-average is

\[
\frac{65 + 95 + 95 + 100}{90 + 100 + 110 + 110} = \frac{355}{410} \approx 86.585\% ,
\]

while the macro-average (averaging them individually) is

\[
\frac{1}{4} \left( \frac{65}{90} + \frac{95}{100} + \frac{95}{110} + \frac{100}{110} \right) \approx 86.124\% .
\]

Because the micro-average is higher, the assignment score contribution would be 86.585%. (The four example scores above do not indicate the total value of the assignments.)

**Grading Scale:** The following grading scale is used on the normalized final, rounded percentages:

If you get at least a/an... and you are in 473, ...and you are in 673, you are guaranteed a/an... or higher.

<table>
<thead>
<tr>
<th>If you get at least...</th>
<th>473—Undergrad</th>
<th>673—Grad</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>A</td>
<td>A-</td>
</tr>
<tr>
<td>80</td>
<td>B</td>
<td>B-</td>
</tr>
<tr>
<td>70</td>
<td>C</td>
<td>C-</td>
</tr>
<tr>
<td>65</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>0</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

3 Dates and Deadlines

Please talk to me if you will have an extended absence.

3.1 Important Dates

Due dates will be announced on the course website and the Piazza discussion board. Unless stated otherwise, items are due by 11:59 PM (UMBC time) of the specified day. Submission instructions will be provided with each assigned item.
3.2 Extensions and Late Policy

Personal or one-off extensions will not be granted. Instead, everyone in this course has ten (10) late days to use as needed throughout the course. These are meant for personal reasons and emergencies; do not use them as an excuse to procrastinate. Late days are measured in 24 hour blocks after a deadline. They are not fractional: an assignment turned in between 1 minute and 23 hours, 59 minutes (1,439 minutes) after the deadline uses one late day, an assignment turned in between 24 hours and 47 hours, 59 minutes (2879) after the deadline uses two late days, etc.

The number of late days remaining has no bearing on assignments you turn in by the deadline; they only affect assignments you turn in after the deadline. If you run out of late days and do not turn an assignment in on time, please still complete and turn in the assignments. Though late assignments after late days have been exhausted will be recorded as a 0, they will still be marked and returned to you. Moreover, they could count in your favor in borderline cases. There is a hard cutoff of the final exam block. Late days cannot be used beyond this time. I reserve the right to issue class-wide extensions.

4 Academic Honesty

Summary  I take academic honesty seriously. Do not cheat or deceive. Doing so may result in lost credit, course failure, or suspension. Instances of suspected dishonesty will be handled through the proper administrative procedures.

This course follows the academic honesty policy from the Office of Undergraduate Education, available from http://oue.umbc.edu/ai/ The following is a concise summary of the policies adopted:

By enrolling in this course, each student assumes the responsibilities of an active participant in UMBCs scholarly community in which everyones academic work and behavior are held to the highest standards of honesty. Cheating, fabrication, plagiarism, and helping others to commit these acts are all forms of academic dishonesty, and they are wrong. Academic misconduct could result in disciplinary action that may include, but is not limited to, suspension or dismissal.

Especially for computer science classes, there are generally questions about what is and is not allowed. You are encouraged to discuss the subject matter and assignments with others. The Piazza discussion board provides a great forum for this. However, you may not write or complete assignments for another student; allow another student to write or complete your assignments; pair program; copy someone else’s work; or allow your work to be copied. (This list is not inclusive.)

As part of discussing the assignments, you may plan with other students; be careful when dealing with pseudocode. A good general rule is that if anything is written down when discussing the assignments with others, you must actually implement it separately and you must not look at your discussion notes.

You are free to use online references like Stack Overflow for questions that are not the primary aspect of the course. If, for example, you’re having an issue with unicode in Python, or are getting a weird compilation error, then sites like Stack Overflow are a great resource. Don’t get stuck fighting your tools.
You may generally use external libraries (and even parts of standard libraries), provided what you use does not actually implement what you are directed to implement.

Be sure to properly acknowledge whatever external help—be it from students, third party libraries, or other readings—you receive.

5 Students with Accommodations

The office of Student Disability Services (SDS, [https://sds.umbc.edu](https://sds.umbc.edu)) works to ensure that students can access and take advantage of UMBC’s educational environment, regardless of disability. From the [SDS](https://sds.umbc.edu), UMBC is committed to eliminating discriminatory obstacles that may disadvantage students based on disability. Services for students with disabilities are provided for all students qualified under the Americans with Disabilities Act (ADA) of 1990, the ADAAA of 2009, and Section 504 of the Rehabilitation Act who request and are eligible for accommodations. The Office of Student Disability Services (SDS) is the UMBC department designated to coordinate accommodations that would allow students to have equal access and inclusion in all courses, programs, and activities at the University.

If you have a documented disability and need to request academic accommodations, please refer to the SDS website at [sds.umbc.edu](http://sds.umbc.edu) for registration information and to begin the process, or alternatively you may visit the SDS office in the Math/Psychology Building, Room 212. For questions or concerns, you may contact us through email at disability@umbc.edu or phone (410) 455-2459.

If you require accommodations for this class, make an appointment to meet with me to discuss your SDS-approved accommodations.

6 Inclusion

College can be stressful. Don’t add to that stress by engaging in harassing or hostile behaviors. They are not welcome in the classroom and are completely inappropriate.

More and more organizations are dealing with hostile behavior. One such organization in the NLP community (NAACL, the North American Association for Computational Linguistics) has a good explanation of the behaviors that are not welcome (replace “conference,” “event” or “ACL” with “class”):

[Harassment and hostile behavior include]: speech or behavior that intimidates, creates discomfort, or interferes with a persons participation or opportunity for participation in a conference or an event. We aim for ACL-related activities to be an environment where harassment in any form does not happen, including but not limited to: harassment based on race, gender, religion, age, color, appearance, national origin, ancestry, disability, sexual orientation, or gender identity. Harassment includes degrading verbal comments, deliberate intimidation, stalking, harassing photography or recording,
inappropriate physical contact, and unwelcome sexual attention. The policy is not intended to inhibit challenging scientific debate, but rather to promote it through ensuring that all are welcome to participate in shared spirit of scientific inquiry.

Any student who has experienced sexual harassment or assault, relationship violence, and/or stalking is encouraged to seek support and resources. There are a number of resources available to you.

Please be aware that as instructors, I and other faculty are required per UMBC policy to report disclosures of sexual assault, domestic violence, relationship violence, stalking, and/or gender-based harassment to the University’s Title IX Coordinator. The purpose of these requirements is for the University to inform you of options, supports, and resources. Please note that Maryland law requires that I report all disclosures or suspicions of child abuse or neglect to the Department of Social Service and/or the police.

You can utilize support and resources even if you do not want to take any further action. You will not be forced to file a police report, but please be aware, depending on the nature of the offense, the University may take action. If you need to speak with someone in confidence about an incident, the following are available:

- The Counseling Center: 410-455-2742 (M-F 8:30-5)
- University Health Services: 410-455-2542 (M-F 8:30-5)
- For after-hours emergency consultation, call the police at 410-455-5555

There are additional, non-confidential on-campus resources:

- The Womens Center (available to students of all genders): 410-455-2714 (M-Th 9:30-6, F 9:30-4)
- Title IX Coordinator: 410-455-1606 (9-5)

7 Version Changes

1.1 8/29/19: Updated TA office hour location
1.0 8/27/19: Initial version